

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 9, 2004, 16:14:15 / Search time 22 Seconds
(without alignments)
2413.642 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815
Sequence: 1 MELALCRWGLLALLPPGA.....TFKGTPTANPEYLGLDPV 1255

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents, AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	6815	100.0	1255 2	US-08-625-101-2 Sequence 2, Appl
2	6815	100.0	1255 2	US-08-356-786-2 Sequence 2, Appl
3	6806	99.9	1255 1	US-08-467-083-68 Sequence 68, Appl
4	6806	99.9	1255 1	US-08-414-417B-68 Sequence 68, Appl
5	6806	99.9	1255 2	US-08-486-348A-68 Sequence 68, Appl
6	6806	99.9	1255 2	US-08-468-545B-68 Sequence 68, Appl
7	6806	99.9	1255 3	US-08-466-680B-68 Sequence 68, Appl
8	6806	99.9	1255 4	US-09-527-487-2 Sequence 2, Appl
9	6732	98.8	1255 2	US-08-484-438-8 Sequence 8, Appl
10	3630	53.3	782 2	US-09-146-283-4 Sequence 8, Appl
11	3630	53.3	782 3	US-08-579-823A-4 Sequence 4, Appl
12	3630	53.3	782 3	US-09-344-195-4 Sequence 4, Appl
13	3473	51.0	624 3	US-08-422-108-1 Sequence 1, Appl
14	3473	51.0	624 4	US-08-422-734-1 Sequence 1, Appl
15	3169	46.5	1210 2	US-08-475-035-4 Sequence 4, Appl
16	3169	46.5	1210 2	US-08-484-438-7 Sequence 7, Appl
17	3091	45.4	580 1	US-08-414-417B-69 Sequence 69, Appl
18	3091	45.4	580 2	US-08-486-348A-69 Sequence 69, Appl
19	3091	45.4	580 3	US-08-468-545B-69 Sequence 69, Appl
20	3091	45.4	580 3	US-08-466-680B-69 Sequence 69, Appl
21	3004.5	44.1	1308 2	US-08-484-438-2 Sequence 2, Appl
22	2871.5	42.1	1058 2	US-08-484-438-4 Sequence 4, Appl
23	2871.5	35.8	1342 1	US-07-978-895-4 Sequence 4, Appl
24	2441.5	35.8	1342 2	US-08-484-438-9 Sequence 9, Appl
25	2441.5	35.8	1342 2	US-08-473-119-4 Sequence 4, Appl
26	2441.5	35.8	1342 2	US-08-475-352-4 Sequence 4, Appl
27	2434	35.7	1343 6	US-08-864-641B-18 Patent No. 5183884

28	1878	27.6	419 4	US-09-630-155-2 Sequence 2, Appl
29	1646.5	24.2	705 2	US-08-456-647B-4 Sequence 4, Appl
30	1646.5	24.2	705 2	US-08-237-401A-9 Sequence 4, Appl
31	1532.5	22.5	644 1	US-08-336-708A-9 Sequence 9, Appl
32	1451	21.3	911 2	US-08-484-438-10 Sequence 10, Appl
33	1389.5	20.4	265 2	US-07-857-224B-66 Sequence 66, Appl
34	1235	18.1	279 2	US-08-701-191A-14 Sequence 14, Appl
35	1188	17.4	541 2	US-08-484-438-6 Sequence 6, Appl
36	1165.5	17.1	265 2	US-07-857-224B-65 Sequence 65, Appl
37	1028.5	15.1	478 4	US-09-570-454-2 Sequence 2, Appl
38	1028.5	15.1	478 4	US-09-867-521-2 Sequence 2, Appl
39	820	12.0	264 2	US-07-857-224B-67 Sequence 67, Appl
40	715	10.5	1382 4	US-08-737-715-2 Sequence 2, Appl
41	715	10.5	1382 4	US-09-457-040B-7 Sequence 7, Appl
42	653	9.6	1367 2	US-08-625-819-2 Sequence 2, Appl
43	651	9.6	1367 2	US-08-249-687C-2 Sequence 2, Appl
44	651	9.6	1367 3	US-08-746-559A-2 Sequence 2, Appl
45	641	9.4	1367 4	US-08-864-641B-18 Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-625-101-2
Sequence 2, Application US/08625101
Patent No. 5868445
GENERAL INFORMATION:
APPLICANT: Cnoveer, Martin A.
APPLICANT: Disib, Mary L.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-625-101-2
Query Match 100.0%; Score 6815; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MELALCRWGLLALLPPGAATGCTGTDKRLPASPETHLDMRLHLYOCQOVQGNL 60
Db 1 MELALCRWGLLALLPPGAATGCTGTDKRLPASPETHLDMRLHLYOCQOVQGNL 60

QY 61 ELTYLPTNASTSLFLODIQEVQGYVLI.AHNQVROVPLORLRI.VRGTOU.FEDNYALAVLDNG 120
 DB 61 ELTYLPTNASTSLFLODIQEVQGYVLI.AHNQVROVPLORLRI.VRGTOU.FEDNYALAVLDNG 120
 QY 121 DELNNTTPVTGASPGGLRELOLSLREILKGVLIQONPOLCYODTILMKDIFHKNOOLA 180
 DB 121 DELNNTTPVTGASPGGLRELOLSLREILKGVLIQONPOLCYODTILMKDIFHKNOOLA 180
 QY 181 LTLIDTNRSPACPCSPMCKGSRCKWGBSSSDCOSLRTTVACAGGACRCKGPLPTDCCHQC 240
 DB 181 LTLIDTNRSPACPCSPMCKGSRCKWGBSSSDCOSLRTTVACAGGACRCKGPLPTDCCHQC 240
 QY 241 AACCTGPXSHDCLAELFHNSGICELHCPALVVTYNTDPESMPNBSGRYTFGASCVTACP 300
 DB 241 AACCTGPXSHDCLAELFHNSGICELHCPALVVTYNTDPESMPNBSGRYTFGASCVTACP 300
 QY 301 YNLTSTDVGSCTVCP.LAHNOBVTABDGTORCEKSKPCAVCYGLGMEHLREVAATVSAN 360
 DB 301 YNLTSTDVGSCTVCP.LAHNOBVTABDGTORCEKSKPCAVCYGLGMEHLREVAATVSAN 360
 QY 361 IQEPACCKIKFGLAFPLPESFDGPASNTAPLOPBOLOVEETLEITGYLIISAMPDLP 420
 DB 361 IQEPACCKIKFGLAFPLPESFDGPASNTAPLOPBOLOVEETLEITGYLIISAMPDLP 420
 QY 421 DLSVFNOLQVIRGRILHNGAYSLTLQGLISWLGSLRELSGGLALIHNTHLCPVHTV 480
 DB 421 DLSVFNOLQVIRGRILHNGAYSLTLQGLISWLGSLRELSGGLALIHNTHLCPVHTV 480
 QY 481 PMOOLPRNHQALHTNRPEDDEVGBGLACHOLCAHGCMGPRPTQVNCOSOTLRQEC 540
 DB 481 PMOOLPRNHQALHTNRPEDDEVGBGLACHOLCAHGCMGPRPTQVNCOSOTLRQEC 540
 QY 541 VEECRVLQGLPREVYVNRHCLPCHPECOPONGSVTCFEPADQCVACAHKDPFCVARC 600
 DB 541 VEECRVLQGLPREVYVNRHCLPCHPECOPONGSVTCFEPADQCVACAHKDPFCVARC 600
 QY 601 PSQVKDPLSTYMP1WKPEDEBGAQCPINCTHSCVDLIDKGCAPRQASPLTISAVG 660
 DB 601 PSQVKDPLSTYMP1WKPEDEBGAQCPINCTHSCVDLIDKGCAPRQASPLTISAVG 660
 QY 720 ILLVVLGVVPGILIKRROOKIRYTMRLLOETELVEPLTPSGAMNOMRLKETEL 720
 DB 720 ILLVVLGVVPGILIKRROOKIRYTMRLLOETELVEPLTPSGAMNOMRLKETEL 720
 QY 780 RYKVLGSGAGFYVYKGIWI.PDGENVKI.PVALIKVRENTSPKANKELIDRAYVMAVGSP 780
 DB 780 RYKVLGSGAGFYVYKGIWI.PDGENVKI.PVALIKVRENTSPKANKELIDRAYVMAVGSP 780
 QY 840 YVSRLLGICLTSYVQVLTOLMPYGCLL.DHVENRGRLSODLNMCMQIAKMSYLEDVR 840
 DB 840 YVSRLLGICLTSYVQVLTOLMPYGCLL.DHVENRGRLSODLNMCMQIAKMSYLEDVR 840
 QY 900 LVRDLAARVLYKSPHNVKITDPGLARLLDIDETETHADGKPIPKMALESTLRRT 900
 DB 900 LVRDLAARVLYKSPHNVKITDPGLARLLDIDETETHADGKPIPKMALESTLRRT 900
 QY 960 HOSDWSYGVYTWELMTFGAKPYDGI.PAREIPDLLEKGERLPOPPICITIDVYMMVMCM 960
 DB 960 HOSDWSYGVYTWELMTFGAKPYDGI.PAREIPDLLEKGERLPOPPICITIDVYMMVMCM 960
 QY 1020 IDSECRPRELVESEFARMADPORFVYI.ONEDLGPASPLDSTFYRSLLDEDDMGDLVDA 1020
 DB 1020 IDSECRPRELVESEFARMADPORFVYI.ONEDLGPASPLDSTFYRSLLDEDDMGDLVDA 1020
 QY 1080 BEYLVPOQGFPCPDPAQAGMVRHRSSTRSGGDLITGLPSEBEARSLASSEG 1080
 DB 1080 BEYLVPOQGFPCPDPAQAGMVRHRSSTRSGGDLITGLPSEBEARSLASSEG 1080
 QY 1140 AGSDVPGDGLGMAAKGLQSLPTHDPSPLQRYSDPTVPLPSERTDGYVAPLTCSPQERY 1140
 DB 1140 AGSDVPGDGLGMAAKGLQSLPTHDPSPLQRYSDPTVPLPSERTDGYVAPLTCSPQERY 1140

QY 1141 NQDVAREPPSPREGELPAPAPAGATLERPKTISBKNGVVKDVAFPGAVENPEVLTPO 1200
 DB 1141 NQDVAREPPSPREGELPAPAPAGATLERPKTISBKNGVVKDVAFPGAVENPEVLTPO 1200
 QY 1201 GGAAPQHPPPAFSPAFNDLYYWDODPPRGAAPPSTFKGPTLANEPEYGLDVPV 1255
 DB 1201 GGAAPQHPPPAFSPAFNDLYYWDODPPRGAAPPSTFKGPTLANEPEYGLDVPV 1255
 RESULT 2
 US-08-356-786-2
 ; Sequence 2, Application US/08356786
 ; Patent No. 5877305
 ; GENERAL INFORMATION:
 ; APPLICANT: Huston, James S.
 ; APPLICANT: Oppermann, Hermann
 ; APPLICANT: Houston, L. L.
 ; APPLICANT: Ring, David B.
 ; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
 ; STREET: Exchange Place, 53 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/356,786
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/831,967
 ; FILING DATE: 06-FEB-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pitcher, Edmund R.
 ; REGISTRATION NUMBER: 27,829
 ; REFERENCE/DOCKET NUMBER: CRP-053
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 248-7000
 ; TELEFAX: (617) 248-7100
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1255 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-356-786-2
 Query Match 100.0%; Score 6815; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELAALCRWMLLALLPFGAASVYCTGTPMKRLPASPTHLDMHLVQGCVOGNTL 60
 DB 1 MELAALCRWMLLALLPFGAASVYCTGTPMKRLPASPTHLDMHLVQGCVOGNTL 60
 QY 61 ELTYLPTNASTSLFLODIQEVQGYVLI.AHNQVROVPLORLRI.VRGTOU.FEDNYALAVLDNG 120
 DB 61 ELTYLPTNASTSLFLODIQEVQGYVLI.AHNQVROVPLORLRI.VRGTOU.FEDNYALAVLDNG 120
 QY 121 DELNNTTPVTGASPGGLRELOLSLREILKGVLIQONPOLCYODTILMKDIFHKNOOLA 180
 DB 121 DELNNTTPVTGASPGGLRELOLSLREILKGVLIQONPOLCYODTILMKDIFHKNOOLA 180
 QY 121 LTLIDTNRSPACPCSPMCKGSRCKWGBSSSDCOSLRTTVACAGGACRCKGPLPTDCCHQC 240
 DB 121 LTLIDTNRSPACPCSPMCKGSRCKWGBSSSDCOSLRTTVACAGGACRCKGPLPTDCCHQC 240

Db 181 LTIIDTNRSRACHPCSPMKSGRCWGBSSBDCOSLTRITVCAAGGCAKCKPPLPTDCHEOC 240
 QY 241 AAGCTGPKHSDCLACHFNHSGICEIHCAPALVTYNTDTFESMNPREGRYTFGASCVTACP 300
 Db 241 AAGCTGPKHSDCLACHFNHSGICEIHCAPALVTYNTDTFESMNPREGRYTFGASCVTACP 300
 QY 301 YNTLSTDVGSCTTVCPHNOEVTABDGTORCEKCSKPCARVCTGAGMEHLREVAVTSAN 360
 Db 301 YNTLSTDVGSCTTVCPHNOEVTABDGTORCEKCSKPCARVCTGAGMEHLREVAVTSAN 360
 QY 361 IOEPACCKIFGSLAFPLESPFDGDPASNTAPLOPELOVETLEETGLYISAMPDILP 420
 Db 361 IOEPACCKIFGSLAFPLESPFDGDPASNTAPLOPELOVETLEETGLYISAMPDILP 420
 QY 421 DLSTVFQNTQVIRGRILHNGAYSLTQGLGIBWLGRLSRLBLSGLALIHNNHLCFVHTV 480
 Db 421 DLSTVFQNTQVIRGRILHNGAYSLTQGLGIBWLGRLSRLBLSGLALIHNNHLCFVHTV 480
 QY 481 PMDOLFRPHQALLHTANRPEBECVGBGLACHQLCARHGCWGPPTQCVNCSQFLRGQBC 540
 Db 481 PMDOLFRPHQALLHTANRPEBECVGBGLACHQLCARHGCWGPPTQCVNCSQFLRGQBC 540
 QY 541 VEEGRVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCAVCAHYXDPPFCVARC 600
 Db 541 VEEGRVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCAVCAHYXDPPFCVARC 600
 QY 601 PSQVXPDLSTYMPIMKFPDEBEGACOPCPINCTHSCVDLDKCGPABQASPLTISIAYVG 660
 Db 601 PSQVXPDLSTYMPIMKFPDEBEGACOPCPINCTHSCVDLDKCGPABQASPLTISIAYVG 660
 QY 661 ILVVVVLGVVGGILIKRQOKIRKTKWRLLOETELVEPLTPSGAMPNOAQMRILKETEL 720
 Db 661 ILVVVVLGVVGGILIKRQOKIRKTKWRLLOETELVEPLTPSGAMPNOAQMRILKETEL 720
 QY 721 RKVAVLGSAGAGTYKGIWIPDGNVVKIPVAIKLRENTSKRAMEILDEAYVMAVGSP 780
 Db 721 RKVAVLGSAGAGTYKGIWIPDGNVVKIPVAIKLRENTSKRAMEILDEAYVMAVGSP 780
 QY 781 YVSRLLGICLSTVQLVTLMPYGLDHYVENRGLSGQDLNMCQIAGMSTLEBYR 840
 Db 781 YVSRLLGICLSTVQLVTLMPYGLDHYVENRGLSGQDLNMCQIAGMSTLEBYR 840
 QY 841 LVHRDLAARNVLVSPNHVKITDFGLARLDIDETBYHADGKVPKIMMALESTLRRFT 900
 Db 841 LVHRDLAARNVLVSPNHVKITDFGLARLDIDETBYHADGKVPKIMMALESTLRRFT 900
 QY 901 HOSDVMVSVTVWMLMTFGAEPYDGIIPAREIPDLLEKGERLPOPPICITIDVYIMVCKM 960
 Db 901 HOSDVMVSVTVWMLMTFGAEPYDGIIPAREIPDLLEKGERLPOPPICITIDVYIMVCKM 960
 QY 961 IDSECRPRELVESEFRMARDPORFVYIOWEDIGPASPILSTFYRSLLEDDMDGLVDA 1020
 Db 961 IDSECRPRELVESEFRMARDPORFVYIOWEDIGPASPILSTFYRSLLEDDMDGLVDA 1020
 QY 1021 EBYLVPOQGFPCPDPAAGAGVHRRHRSSTRSGGDLTLGLEPSEBEARSPILASEG 1080
 Db 1021 EBYLVPOQGFPCPDPAAGAGVHRRHRSSTRSGGDLTLGLEPSEBEARSPILASEG 1080
 QY 1081 AGSDVPRGDLGMAKAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQREYV 1140
 Db 1081 AGSDVPRGDLGMAKAGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQREYV 1140
 QY 1141 NQPDVPRQPSPREGLPAAAPAGATLERPKTILSGKNGVYKDVAFAGGAVENREYVLPQ 1200
 Db 1141 NQPDVPRQPSPREGLPAAAPAGATLERPKTILSGKNGVYKDVAFAGGAVENREYVLPQ 1200
 QY 1201 GGAAPRHPAPSPAFDNTLYMDODPBERGAPSTFKGTPTAENPYLGLDVPV 1255
 Db 1201 GGAAPRHPAPSPAFDNTLYMDODPBERGAPSTFKGTPTAENPYLGLDVPV 1255

RESULT 3
 US-08-467-083-68

; Sequence 68, Application US/08467083
 ; Patent No. 5726023
 ; GENERAL INFORMATION:
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: Disib, Mary L.
 ; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
 ; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
 ; NUMBER OF SEQUENCES: 68
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Seed and Berry
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: US
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/467,083
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/414,417
 ; FILING DATE: 06-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sharkey, Richard G.
 ; REGISTRATION NUMBER: 32,629
 ; REFERENCE/DOCKET NUMBER: 920010.448C2
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; TELEX: 3723836 SEDANBERY
 ; INFORMATION FOR SEQ ID NO: 68:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1255 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-467-083-68
 Query Match 99.9%; Score 6806; DB 1; Length 1255;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 METALACRMGILLALLPFGAASVQCTGTDKRLPASPEHLMRLHYOGCOVVQGNL 60
 Db 1 METALACRMGILLALLPFGAASVQCTGTDKRLPASPEHLMRLHYOGCOVVQGNL 60
 QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIANOVQVPLQRLRIVRSTQJFEDNYALAVLDNG 120
 Db 61 ELTYLPTNASLSFLQDIOEVQGYVLIANOVQVPLQRLRIVRSTQJFEDNYALAVLDNG 120
 QY 121 DPLNNTPTVYGAAPGGRRLQRLSLTEILKGVLIQNNPOLCYDDITLMKDI FHKNNOLA 180
 Db 121 DPLNNTPTVYGAAPGGRRLQRLSLTEILKGVLIQNNPOLCYDDITLMKDI FHKNNOLA 180
 QY 181 LTLIDTNRSRACHPCSPMKSGRCWGBSSBDCOSLTRITVCAAGGCAKCKPPLPTDCHEOC 240
 Db 181 LTLIDTNRSRACHPCSPMKSGRCWGBSSBDCOSLTRITVCAAGGCAKCKPPLPTDCHEOC 240
 QY 241 AAGCTGPKHSDCLACHFNHSGICEIHCAPALVTYNTDTFESMNPREGRYTFGASCVTACP 300
 Db 241 AAGCTGPKHSDCLACHFNHSGICEIHCAPALVTYNTDTFESMNPREGRYTFGASCVTACP 300
 QY 301 YNTLSTDVGSCTTVCPHNOEVTABDGTORCEKCSKPCARVCTGAGMEHLREVAVTSAN 360
 Db 301 YNTLSTDVGSCTTVCPHNOEVTABDGTORCEKCSKPCARVCTGAGMEHLREVAVTSAN 360
 QY 361 IOEPACCKIFGSLAFPLESPFDGDPASNTAPLOPELOVETLEETGLYISAMPDILP 420

Db 361 IGFACCKKI FGSGLAFPSFDGDPASNTAPLOPBOLOVEETLEETGYLYISAMPDLSL 420
 Qy 421 DLSVFONLOVIRIRILHNGAYSLTLOGIGISWGLASRLBSGLALIHNTLCEVHTV 480
 Db 421 DLSVFONLOVIRIRILHNGAYSLTLOGIGISWGLASRLBSGLALIHNTLCEVHTV 480
 Qy 481 PMDOLFRNPHOALLHTANRPEDECEVEGLACHOLCARHCWGPGPTQCVNCSQFLRGQEC 540
 Db 481 PMDOLFRNPHOALLHTANRPEDECEVEGLACHOLCARHCWGPGPTQCVNCSQFLRGQEC 540
 Qy 541 VEECRVLOGLPREYVNAHCLPCHPCOPONGSVTCFPGPADOCVACAHYKDPFCVARC 600
 Db 541 VEECRVLOGLPREYVNAHCLPCHPCOPONGSVTCFPGPADOCVACAHYKDPFCVARC 600
 Qy 601 PGVPRDLSTYMPKRPDEBGACQPCPINCSTHCVLDKGCARASPLTISIAVVG 660
 Db 601 PGVPRDLSTYMPKRPDEBGACQPCPINCSTHCVLDKGCARASPLTISIAVVG 660
 Qy 661 ILLVVVLGVVFGILLRROOKIRKTYTRRLLOETELVEPLTPSGAMPNOQMRILKETE 720
 Db 661 ILLVVVLGVVFGILLRROOKIRKTYTRRLLOETELVEPLTPSGAMPNOQMRILKETE 720
 Qy 721 RYKTVLGSAGFVYVGIWIPDEENKIPALVYLAENTSPPRANKIILDEAYYMAVGS 780
 Db 721 RYKTVLGSAGFVYVGIWIPDEENKIPALVYLAENTSPPRANKIILDEAYYMAVGS 780
 Qy 781 YSRLLGICLTSTVOLTOLMPYGCLLDHYBNRGLSGODLLNMCQIAGMSYLEDVR 840
 Db 781 YSRLLGICLTSTVOLTOLMPYGCLLDHYBNRGLSGODLLNMCQIAGMSYLEDVR 840
 Qy 841 LVHRDLAARNVLYKSPNHVYKTDGFLARLLDIDETRYHADGKVPKMMALLESILRRPT 900
 Db 841 LVHRDLAARNVLYKSPNHVYKTDGFLARLLDIDETRYHADGKVPKMMALLESILRRPT 900
 Qy 901 HOSDWVSXGVTWELMTFGAKPYDGI PARI PDLLEKGERLPOPPCTIDVYIMKCM 960
 Db 901 HOSDWVSXGVTWELMTFGAKPYDGI PARI PDLLEKGERLPOPPCTIDVYIMKCM 960
 Qy 961 IOSECRPRELVESESRMARDPORFVIONEDLGPASPLDSTFYSLIEDDMGDLVDA 1020
 Db 961 IOSECRPRELVESESRMARDPORFVIONEDLGPASPLDSTFYSLIEDDMGDLVDA 1020
 Qy 1021 BEYLVOQGFPCDPAAGAGVNHRRSSSTRSGGGLTLGLPSEBEAPRPLAPSEG 1080
 Db 1021 BEYLVOQGFPCDPAAGAGVNHRRSSSTRSGGGLTLGLPSEBEAPRPLAPSEG 1080
 Qy 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLORYSEDPVPLPSETDGVAPLTCSPQREYV 1140
 Db 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLORYSEDPVPLPSETDGVAPLTCSPQREYV 1140
 Qy 1141 NOPDVAPPPRPPRPLPARPAGATLERPKTSPGKNGVYKDVAFAGAVENPEYLPQ 1200
 Db 1141 NOPDVAPPPRPPRPLPARPAGATLERPKTSPGKNGVYKDVAFAGAVENPEYLPQ 1200
 Qy 1201 GGAAPRPHRPPASPAFDNLVYWDOPRERGAPESTFKGTPTAENPEYLGDLV 1255
 Db 1201 GGAAPRPHRPPASPAFDNLVYWDOPRERGAPESTFKGTPTAENPEYLGDLV 1255

RESULT 4
 US-08-414-417B-68
 Sequence 68, Application US/08414417B
 Patent No. 5801005
 GENERAL INFORMATION:
 APPLICANT: Cheever, Martin A.
 APPLICANT: Diste, Mary L.
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
 TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
 NUMBER OF SEQUENCES: 69
 CORRESPONDENCE ADDRESSES:
 ADDRESSES: Seed and Betty LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
 STATE: Washington
 COUNTRY: US
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/414,417B
 FILING DATE: 31-MAR-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Sharkey, Richard G.
 REGISTRATION NUMBER: 32,629
 REFERENCE/DOCKET NUMBER: 920010.448C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 68:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1255 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-414-417B-68

Query Match 99.9%; Score 6806; DB 1; Length 1255;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 METALCEWGLLALLPAGASTOVCTGDMKLRLEPASETHIDMLRHLVYOGCQVVOG 60
 Db 1 METALCEWGLLALLPAGASTOVCTGDMKLRLEPASETHIDMLRHLVYOGCQVVOG 60
 Qy 61 ELTYLPTNASTSFLDIOEVQGVYLAHQVQVPLRRLRYVGTOLFEDNYALAVLDNG 120
 Db 61 ELTYLPTNASTSFLDIOEVQGVYLAHQVQVPLRRLRYVGTOLFEDNYALAVLDNG 120
 Qy 121 DEANNTPTVGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILWMDIFHKNNOLA 180
 Db 121 DEANNTPTVGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILWMDIFHKNNOLA 180
 Qy 181 LTLIDTNRACHPCSPMKSGSRCEWSESSDCQSLTRTYCAGGCANCKGFLPTDCHEOC 240
 Db 181 LTLIDTNRACHPCSPMKSGSRCEWSESSDCQSLTRTYCAGGCANCKGFLPTDCHEOC 240
 Qy 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTESMNPBGRVTFEGASCVTACP 300
 Db 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTESMNPBGRVTFEGASCVTACP 300
 Qy 301 YNLTSTDVSCITVCLHNGEYTAEDGTORCEKSRPCARVCYGLGMEHLREYAVTSAN 360
 Db 301 YNLTSTDVSCITVCLHNGEYTAEDGTORCEKSRPCARVCYGLGMEHLREYAVTSAN 360
 Qy 361 IGFACCKKI FGSGLAFPSFDGDPASNTAPLOPBOLOVEETLEETGYLYISAMPDLSL 420
 Db 361 IGFACCKKI FGSGLAFPSFDGDPASNTAPLOPBOLOVEETLEETGYLYISAMPDLSL 420
 Qy 421 DLSVFONLOVIRIRILHNGAYSLTLOGIGISWGLASRLBSGLALIHNTLCEVHTV 480
 Db 421 DLSVFONLOVIRIRILHNGAYSLTLOGIGISWGLASRLBSGLALIHNTLCEVHTV 480
 Qy 481 PMDOLFRNPHOALLHTANRPEDECEVEGLACHOLCARHCWGPGPTQCVNCSQFLRGQEC 540
 Db 481 PMDOLFRNPHOALLHTANRPEDECEVEGLACHOLCARHCWGPGPTQCVNCSQFLRGQEC 540
 Qy 541 VEECRVLOGLPREYVNAHCLPCHPCOPONGSVTCFPGPADOCVACAHYKDPFCVARC 600
 Db 541 VEECRVLOGLPREYVNAHCLPCHPCOPONGSVTCFPGPADOCVACAHYKDPFCVARC 600
 Qy 601 PGVPRDLSTYMPKRPDEBGACQPCPINCSTHCVLDKGCARASPLTISIAVVG 660

```

Db 601 PSQVPLDLYMPEIKWPEDEBGAQCPICNCHTSVDLDKGCAPBQAPLTIISAVVG 660
Qy 661 ILVVVLGVVFGIILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOAMRIKETEL 720
Db 661 ILVVVLGVVFGIILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOAMRIKETEL 720
Qy 721 RKVKVLGSGAFGTYYKGIWIPDGENVKI PVAIKYLRRENTSPKANKIIDEAYVAVG 780
Db 721 RKVKVLGSGAFGTYYKGIWIPDGENVKI PVAIKYLRRENTSPKANKIIDEAYVAVG 780
Qy 781 YVSRLLGICLTSTVOLVTOAMPYGLLDHYRENGRLSGODLNMCMQIAKMSYLEDR 840
Db 781 YVSRLLGICLTSTVOLVTOAMPYGLLDHYRENGRLSGODLNMCMQIAKMSYLEDR 840
Qy 841 LVHRDLAARNVLVKS PNHVKITDFGLARLLDIDETEHADGSKVPIKMMALSIIRRT 900
Db 841 LVHRDLAARNVLVKS PNHVKITDFGLARLLDIDETEHADGSKVPIKMMALSIIRRT 900
Qy 901 HOSDVMSYGVYVWELMTFGAKPYDGI PAKEIPDLLEKGERLPQPICTIDVYMIWKCMM 960
Db 901 HOSDVMSYGVYVWELMTFGAKPYDGI PAKEIPDLLEKGERLPQPICTIDVYMIWKCMM 960
Qy 961 IDSECRPFRELVSFESRMARDPQRFVYVIONEDLPASPLDSTFYRSLLEDMDGLDVA 1020
Db 961 IDSECRPFRELVSFESRMARDPQRFVYVIONEDLPASPLDSTFYRSLLEDMDGLDVA 1020
Qy 1021 EEYLVPQGFPCPDPAFGAGVWHRHRSSTRSGGDLTLGLBSEBEPAPSPAPSEB 1080
Db 1021 EEYLVPQGFPCPDPAFGAGVWHRHRSSTRSGGDLTLGLBSEBEPAPSPAPSEB 1080
Qy 1081 AGSDVFPDDLMGAKGLQSLPTHDPSPLOKYSBDPTVPLPSETGYVAPLTCSPQEV 1140
Db 1081 AGSDVFPDDLMGAKGLQSLPTHDPSPLOKYSBDPTVPLPSETGYVAPLTCSPQEV 1140
Qy 1141 NOPDVRPQPPREBEPAPARPAAGATLERPKTLSPKGVVADVAPGAVENPEYLPQ 1200
Db 1141 NOPDVRPQPPREBEPAPARPAAGATLERPKTLSPKGVVADVAPGAVENPEYLPQ 1200
Qy 1201 GGAAPQPPAPPAFPADNLVYMDPPERGAPEPTFAENPEYLGIDVPY 1255
Db 1201 GGAAPQPPAPPAFPADNLVYMDPPERGAPEPTFAENPEYLGIDVPY 1255

```

RESULT 5

US-08-486-348A-68
Sequence 68, Application US/08486348A
Patent No. 5846538

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.
APPLICANT: Dials, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSER: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629

```

REFERENCE/DOCKET NUMBER: 920010.448C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match: 99.9%; Score 6806; DB 2; Length 1255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 METALALCWGILLALLPPGAASVYCTGTDMKRLPASEFTHLDMRLHYOGGVVQNL 60
1 METALALCWGILLALLPPGAASVYCTGTDMKRLPASEFTHLDMRLHYOGGVVQNL 60
61 ELTYLPTNASLFLQDIOEVQGYVLIAHNOYRQVPLQRLRIYRGTQLFEDNTALAVLDNG 120
61 ELTYLPTNASLFLQDIOEVQGYVLIAHNOYRQVPLQRLRIYRGTQLFEDNTALAVLDNG 120
121 PLNNTTPTVGAAPGELRELOLRSLTEILKGVLIQRNPOLCYODTIIMKDI FHKNNOLA 180
121 PLNNTTPTVGAAPGELRELOLRSLTEILKGVLIQRNPOLCYODTIIMKDI FHKNNOLA 180
181 LTIIDTNSRACHPSPCKSKSRCKGSESDQSLTRTYCAGGACRCKPLPTDCHEQC 240
181 LTIIDTNSRACHPSPCKSKSRCKGSESDQSLTRTYCAGGACRCKPLPTDCHEQC 240
241 AAGCTGPGHSDCLACIHNHSGICELHCPALVTYVTFPESMPNEGRTPGASCTYAP 300
241 AAGCTGPGHSDCLACIHNHSGICELHCPALVTYVTFPESMPNEGRTPGASCTYAP 300
301 YNLTSTVGSCTIVCPHNOEVTABDGTORCEKSKPCARVYCGMEHLREVRATVSN 360
301 YNLTSTVGSCTIVCPHNOEVTABDGTORCEKSKPCARVYCGMEHLREVRATVSN 360
361 IOEPAGCKKIGSLAFLESTFDGPASNTAPLOPQLOVFETLEITGYLYISAMPDSL 420
361 IOEPAGCKKIGSLAFLESTFDGPASNTAPLOPQLOVFETLEITGYLYISAMPDSL 420
421 DLSTFONLQVIRGLIHNHGYSLTQGLISWLGIRSLREISGSLALHNHNLCPVHTV 480
421 DLSTFONLQVIRGLIHNHGYSLTQGLISWLGIRSLREISGSLALHNHNLCPVHTV 480
481 PMDQLFNPQALHTANRPEDECVGEGLACHOLCARGHCWGPPTOCVNSQFLRGQEC 540
481 PMDQLFNPQALHTANRPEDECVGEGLACHOLCARGHCWGPPTOCVNSQFLRGQEC 540
541 VESCRVIOGLREYVNAHCLPCHPECOPONGSVTCFEPBADQCYAACHYDPPCVAR 600
541 VESCRVIOGLREYVNAHCLPCHPECOPONGSVTCFEPBADQCYAACHYDPPCVAR 600
601 PSQVPLDLYMPEIKWPEDEBGAQCPICNCHTSVDLDKGCAPBQAPLTIISAVVG 660
601 PSQVPLDLYMPEIKWPEDEBGAQCPICNCHTSVDLDKGCAPBQAPLTIISAVVG 660
661 ILVVVLGVVFGIILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOAMRIKETEL 720
661 ILVVVLGVVFGIILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOAMRIKETEL 720
721 RKVKVLGSGAFGTYYKGIWIPDGENVKI PVAIKYLRRENTSPKANKIIDEAYVAVG 780
721 RKVKVLGSGAFGTYYKGIWIPDGENVKI PVAIKYLRRENTSPKANKIIDEAYVAVG 780
781 YVSRLLGICLTSTVOLVTOAMPYGLLDHYRENGRLSGODLNMCMQIAKMSYLEDR 840
781 YVSRLLGICLTSTVOLVTOAMPYGLLDHYRENGRLSGODLNMCMQIAKMSYLEDR 840
841 LVHRDLAARNVLVKS PNHVKITDFGLARLLDIDETEHADGSKVPIKMMALSIIRRT 900
841 LVHRDLAARNVLVKS PNHVKITDFGLARLLDIDETEHADGSKVPIKMMALSIIRRT 900

```

Db 841 LVHRDLAARNVLYKSPNHYKTDPGIARLLDIDETETHADGKVPYKMALESILRRFT 900
 Qy 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYMMVKCM 960
 Db 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYMMVKCM 960
 Qy 961 IDSECRPRRELVSERFARMARDPQRFVVIQNEDLGPASPLDSTFYRSLEDDMGDLVDA 1020
 Db 961 IDSECRPRRELVSERFARMARDPQRFVVIQNEDLGPASPLDSTFYRSLEDDMGDLVDA 1020
 Qy 1021 BEYLVPQGFPCDPDPAAGAGWVHHRSSSTRSGGDLTLGLEPSEEBAPSLAPSEG 1080
 Db 1021 BEYLVPQGFPCDPDPAAGAGWVHHRSSSTRSGGDLTLGLEPSEEBAPSLAPSEG 1080
 Qy 1081 AGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPVPLPSETDGYVALTCSPOPEYV 1140
 Db 1081 AGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPVPLPSETDGYVALTCSPOPEYV 1140
 Qy 1141 NOPDVRPQPSRREGPLPARPAGATLERPKTISPGKGVYKDVPAFGAVENPEYLTDQ 1200
 Db 1141 NOPDVRPQPSRREGPLPARPAGATLERPKTISPGKGVYKDVPAFGAVENPEYLTDQ 1200
 Qy 1201 GGAAPQHPHPPAPSPAFDNLVYWDODPPRGAPESTFKGPTAENPEYLGIDVPV 1255
 Db 1201 GGAAPQHPHPPAPSPAFDNLVYWDODPPRGAPESTFKGPTAENPEYLGIDVPV 1255

RESULT 6
 US-08-468-545B-68
 / Sequence 68, Application US/08468545B
 / Patent No. 5876712

/ GENERAL INFORMATION:
 / APPLICANT: Cheever, Martin A.
 / APPLICANT: Delsis, Mary L.
 / TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
 / TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
 / TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
 / NUMBER OF SEQUENCES: 69
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Seed and Berry LLP
 / STREET: 6300 Columbia Center, 701 Fifth Avenue
 / CITY: Seattle
 / STATE: Washington
 / COUNTRY: US
 / ZIP: 98104-7092
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/468,545B
 / FILING DATE: 06-JUN-1995
 / CLASSIFICATION: 424
 / ATTORNEY/AGENT INFORMATION:
 / .NAME: Sharkey, Richard G.
 / REGISTRATION NUMBER: 32,629
 / REFERENCE/DOCKET NUMBER: 920010.448CS
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (206) 622-4900
 / TELEFAX: (206) 682-6031
 / INFORMATION FOR SEQ ID NO: 68:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1255 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / US-08-468-545B-68

Query Match 99.9%; Score 6806; DB 2; Length 1255;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MELAALCRWGLLALLPFGAASGVCTGTDMKRLPASPETHLDMLRHLVQGCQVQGNL 60

Db 1 MELAALCRWGLLALLPFGAASGVCTGTDMKRLPASPETHLDMLRHLVQGCQVQGNL 60
 Qy 61 ELTYLPTNASLSFLDIOEVQGYVLIANQVQVPLQIRLIRYGTQLFEDNVALAVLDNG 120
 Db 61 ELTYLPTNASLSFLDIOEVQGYVLIANQVQVPLQIRLIRYGTQLFEDNVALAVLDNG 120
 Qy 121 DELNNTTPVTGASPGGLBELQLRSITELIKGGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
 Db 121 DELNNTTPVTGASPGGLBELQLRSITELIKGGVLIQRNPOLCYODTILMKDIFHKNNOLA 180
 Qy 181 LTLIDTNSRACHPSPCKSGSRGSESDOSLTRTYCAGGCAKCPPTDCCHEOC 240
 Db 181 LTLIDTNSRACHPSPCKSGSRGSESDOSLTRTYCAGGCAKCPPTDCCHEOC 240
 Qy 241 AAGCTGPKASDCLACIHFHNSGICELHCPALVTYNTDRESMPNPGRYTFGASCYTACP 300
 Db 241 AAGCTGPKASDCLACIHFHNSGICELHCPALVTYNTDRESMPNPGRYTFGASCYTACP 300
 Qy 301 YNLTSTVSGCTLVCPRLNQEVTAEDGTORCRKSCFPCARVCYGLMELHREYRAVTSAN 360
 Db 301 YNLTSTVSGCTLVCPRLNQEVTAEDGTORCRKSCFPCARVCYGLMELHREYRAVTSAN 360
 Qy 361 IGEFAGCKKI FGSILAFLEPSFDGDPASNTAPLOPELOVFEETLEITGYLISAMPDSL 420
 Db 361 IGEFAGCKKI FGSILAFLEPSFDGDPASNTAPLOPELOVFEETLEITGYLISAMPDSL 420
 Qy 421 DLSVFQNLQVIRGRILHNGVSLTLQGLGISWLGRLSRLBLSGLALIHNTLCEVHTV 480
 Db 421 DLSVFQNLQVIRGRILHNGVSLTLQGLGISWLGRLSRLBLSGLALIHNTLCEVHTV 480
 Qy 481 PMDOLFRRNHQALLHTANRPEDBCVSGEGLACHQLCARHCWGSGPQCVCNCSQFLRQEC 540
 Db 481 PMDOLFRRNHQALLHTANRPEDBCVSGEGLACHQLCARHCWGSGPQCVCNCSQFLRQEC 540
 Qy 541 VECRVLQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVCACAHKDPFCVARC 600
 Db 541 VECRVLQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVCACAHKDPFCVARC 600
 Qy 601 PSGVKEPDLSPMPIMKPEDEGACQPCPINCTHSCVDLDKGCAPAEORASPLTISAVNG 660
 Db 601 PSGVKEPDLSPMPIMKPEDEGACQPCPINCTHSCVDLDKGCAPAEORASPLTISAVNG 660
 Qy 661 ILLVVLGVVFGILLIRROOKIRKTYMRLLQETELVEPLTPSGAMPNOAMRLKETEL 720
 Db 661 ILLVVLGVVFGILLIRROOKIRKTYMRLLQETELVEPLTPSGAMPNOAMRLKETEL 720
 Qy 721 RKVKVLGSGAFGVYKGIWIPGSENVKIPALRVLEBNSPKXNKELIDEAYYMAAGVSP 780
 Db 721 RKVKVLGSGAFGVYKGIWIPGSENVKIPALRVLEBNSPKXNKELIDEAYYMAAGVSP 780
 Qy 781 YNSRLIGICLTSTVOLVQLMPYGCILLDHVRENGRLGSDLLNMCQIAKMSYLEDDR 840
 Db 781 YNSRLIGICLTSTVOLVQLMPYGCILLDHVRENGRLGSDLLNMCQIAKMSYLEDDR 840
 Qy 841 LVHRDLAARNVLYKSPNHYKTDPGIARLLDIDETETHADGKVPYKMALESILRRFT 900
 Db 841 LVHRDLAARNVLYKSPNHYKTDPGIARLLDIDETETHADGKVPYKMALESILRRFT 900
 Qy 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYMMVKCM 960
 Db 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYMMVKCM 960
 Qy 961 IDSECRPRRELVSERFARMARDPQRFVVIQNEDLGPASPLDSTFYRSLEDDMGDLVDA 1020
 Db 961 IDSECRPRRELVSERFARMARDPQRFVVIQNEDLGPASPLDSTFYRSLEDDMGDLVDA 1020
 Qy 1021 BEYLVPQGFPCDPDPAAGAGWVHHRSSSTRSGGDLTLGLEPSEEBAPSLAPSEG 1080
 Db 1021 BEYLVPQGFPCDPDPAAGAGWVHHRSSSTRSGGDLTLGLEPSEEBAPSLAPSEG 1080
 Qy 1081 AGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPVPLPSETDGYVALTCSPOPEYV 1140

Db 1081 AGSDVDFDGLGMAKGLQSLPTHDPSPLOKRYSEDPVPLPSSETDGVAVPLTCSPOPEYV 1140
Qy 1141 NOPDVPRPSPSPREGULPAARPAAGATLTERPXTLSPKNGVYKDVAFGAVENPEYLTPQ 1200
Db 1141 NOPDVPRPSPSPREGULPAARPAAGATLTERPXTLSPKNGVYKDVAFGAVENPEYLTPQ 1200
Qy 1201 GGAAPQHPHPAPSPADNLTYMDODPPERGAPSPSTFKGTPTAENPEYLGLDVPV 1255
Db 1201 GGAAPQHPHPAPSPADNLTYMDODPPERGAPSPSTFKGTPTAENPEYLGLDVPV 1255

RESULT 7
US-08-466-6808-68
; Sequence 68, Application US/084666808
; Patent No. 6075122
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Dieig, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,6808
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-466-6808-68

Query Match 99.9%; Score 6806; DB 3; Length 1255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAALCRMGILLALPPGAASVQCTGTDMLRLPASPETHLDMLRLHYOGCOVVOGNTL 60
Db 1 MELAALCRMGILLALPPGAASVQCTGTDMLRLPASPETHLDMLRLHYOGCOVVOGNTL 60
Qy 61 ELTYLPTNASTSPLQDIOEVQGYVLIANOVQVPLQSLRIVRGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASTSPLQDIOEVQGYVLIANOVQVPLQSLRIVRGTQLFEDNYALAVLDNG 120
Qy 121 DPLANNTPVTGASPGSLBELQSLTELTKGVLQORPQCYQDTIMKQIFPKNNOLA 180
Db 121 DPLANNTPVTGASPGSLBELQSLTELTKGVLQORPQCYQDTIMKQIFPKNNOLA 180
Qy 181 LTLIDTRRSRACHPCSPWCKSRCKGSESDCQSLTRVYCGGACRCKGPLPTDCCHQC 240
Db 181 LTLIDTRRSRACHPCSPWCKSRCKGSESDCQSLTRVYCGGACRCKGPLPTDCCHQC 240
Qy 241 AAGCTGRKSHDCLACHFNHSGICEHLCPALVYNTDTFESMPNBERYTFGASCVTACP 300

Db 241 AAGCTGRKSHDCLACHFNHSGICEHLCPALVYNTDTFESMPNBERYTFGASCVTACP 300
Qy 301 YNLTSTDVSCCTLVCCILHNOETABEDGTORCEKSKRCAVCCGLMEHLREVAATVSAN 360
Db 301 YNLTSTDVSCCTLVCCILHNOETABEDGTORCEKSKRCAVCCGLMEHLREVAATVSAN 360
Qy 361 IOEFACCKKIFGSLAEPLSPFODGPASNTAPLQPEOLQVETLEIETGYLYISAWPSLP 420
Db 361 IOEFACCKKIFGSLAEPLSPFODGPASNTAPLQPEOLQVETLEIETGYLYISAWPSLP 420
Qy 421 DISVFQNLQVIRGRILHNGAYSLTLQGLISWLGSLRLRELSGLALIHNTHLCPVHTV 480
Db 421 DISVFQNLQVIRGRILHNGAYSLTLQGLISWLGSLRLRELSGLALIHNTHLCPVHTV 480
Qy 481 PMDOLFRNHQALLHNAHPPEDECVESGLAQHLCARGHCWGSPPTCCVNCOSQFLRQEC 540
Db 481 PMDOLFRNHQALLHNAHPPEDECVESGLAQHLCARGHCWGSPPTCCVNCOSQFLRQEC 540
Qy 541 VEECRVLQGLPREVYNARHCLPCHPECOFONGSVTCFGEADQVCAHAKDBPCVARC 600
Db 541 VEECRVLQGLPREVYNARHCLPCHPECOFONGSVTCFGEADQVCAHAKDBPCVARC 600
Qy 601 PSQVPRDLSYMPIMKFPDEGACOPCPINCTHSCVDLDDKGCAPAEORASPLTSISAVG 660
Db 601 PSQVPRDLSYMPIMKFPDEGACOPCPINCTHSCVDLDDKGCAPAEORASPLTSISAVG 660
Qy 661 ILVYVVLGVVFGILLKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOQMRILKETEL 720
Db 661 ILVYVVLGVVFGILLKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOQMRILKETEL 720
Qy 721 RKVKVLGSGAGSTYKGIWIPGENVVKIPVALKULRENTSPRANKIILDEAVYAGVGP 780
Db 721 RKVKVLGSGAGSTYKGIWIPGENVVKIPVALKULRENTSPRANKIILDEAVYAGVGP 780
Qy 781 YVSRLLGICLTSTVOLVTLQMPYGLLDHVRNRRGLSGQDLLNMCQIAKGSYLEYDR 840
Db 781 YVSRLLGICLTSTVOLVTLQMPYGLLDHVRNRRGLSGQDLLNMCQIAKGSYLEYDR 840
Qy 841 LVHRDLAARNVLYSPNNVKTIDFGLARLLIDETERYHADGKVPKIMALESILRRFT 900
Db 841 LVHRDLAARNVLYSPNNVKTIDFGLARLLIDETERYHADGKVPKIMALESILRRFT 900
Qy 901 HOSDVMSYGVYVWELMTGAKPYDGIIPAREIPDLEKGRLEPOPTCTIDVYIMVCKWM 960
Db 901 HOSDVMSYGVYVWELMTGAKPYDGIIPAREIPDLEKGRLEPOPTCTIDVYIMVCKWM 960
Qy 961 IDSECRPRFRELVSEFSHARDPORFVYIYONEDLGPASPLDSTFYRSLEDDMDGLVDA 1020
Db 961 IDSECRPRFRELVSEFSHARDPORFVYIYONEDLGPASPLDSTFYRSLEDDMDGLVDA 1020
Qy 1021 EETLVPOQGFPCPDPAAGAGGVHRRHSSSTRSGGDLTLGLBPSSEERAPSPAPSEG 1080
Db 1021 EETLVPOQGFPCPDPAAGAGGVHRRHSSSTRSGGDLTLGLBPSSEERAPSPAPSEG 1080
Qy 1081 AGSDVDFDGLGMAKGLQSLPTHDPSPLOKRYSEDPVPLPSSETDGVAVPLTCSPOPEYV 1140
Db 1081 AGSDVDFDGLGMAKGLQSLPTHDPSPLOKRYSEDPVPLPSSETDGVAVPLTCSPOPEYV 1140
Qy 1141 NOPDVPRPSPSPREGULPAARPAAGATLTERPXTLSPKNGVYKDVAFGAVENPEYLTPQ 1200
Db 1141 NOPDVPRPSPSPREGULPAARPAAGATLTERPXTLSPKNGVYKDVAFGAVENPEYLTPQ 1200
Qy 1201 GGAAPQHPHPAPSPADNLTYMDODPPERGAPSPSTFKGTPTAENPEYLGLDVPV 1255
Db 1201 GGAAPQHPHPAPSPADNLTYMDODPPERGAPSPSTFKGTPTAENPEYLGLDVPV 1255

RESULT 8
US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:

APPLICANT: Nicolette, Charles
 TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
 FILE REFERENCE: 126881309200
 CURRENT APPLICATION NUMBER: US/09/527,487
 CURRENT FILING DATE: 2000-03-16
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 2
 LENGTH: 1255
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-527-487-2

Query Match 99.9%; Score 6806; DB 4; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MELAALCRWGLIALALPPGAASVCTGDMKRLPASPEHLDMLRHLYOGGVQGNL 60
 1 MELAALCRWGLIALALPPGAASVCTGDMKRLPASPEHLDMLRHLYOGGVQGNL 60
 61 ELTYLFTNALSIFLODIOEVQVYLIAHQVROVPLQRLIYVGTOLFEEDNALAVLDNG 120
 61 ELTYLFTNALSIFLODIOEVQVYLIAHQVROVPLQRLIYVGTOLFEEDNALAVLDNG 120
 121 DPLANTPTVAGASPGGLRELQRLSLTEILKGVLIQANPOLCYDTITLMDKDIHKNQOLA 180
 121 DPLANTPTVAGASPGGLRELQRLSLTEILKGVLIQANPOLCYDTITLMDKDIHKNQOLA 180
 121 DPLANTPTVAGASPGGLRELQRLSLTEILKGVLIQANPOLCYDTITLMDKDIHKNQOLA 180
 181 LTLIDNRSBACHPCGPMCKGSRCWGSSSEDCOSLRTTVAAGCARGKGPLPTDCHEOC 240
 181 LTLIDNRSBACHPCGPMCKGSRCWGSSSEDCOSLRTTVAAGCARGKGPLPTDCHEOC 240
 241 AAGCTGPKGSDCLACLFHNSGICELHCPALVTNTDTFESMNPBGRYTFGASCVTACP 300
 241 AAGCTGPKGSDCLACLFHNSGICELHCPALVTNTDTFESMNPBGRYTFGASCVTACP 300
 301 YNTLSDVGSCTIYVCPHANOVTABDGTQCEKSCFCAVVCYGLGMEHLREYRAVTSAN 360
 301 YNTLSDVGSCTIYVCPHANOVTABDGTQCEKSCFCAVVCYGLGMEHLREYRAVTSAN 360
 361 IOEFACCKIKFGSLAPFESFDDPASNTPALQPEOLQVETLEITGYIISAMPDLP 420
 361 IOEFACCKIKFGSLAPFESFDDPASNTPALQPEOLQVETLEITGYIISAMPDLP 420
 421 DLSVPONLQVIRGRILHNGAYSLTLQGLISWLGRLRELSGGLAIHNTHLCFVHTV 480
 421 DLSVPONLQVIRGRILHNGAYSLTLQGLISWLGRLRELSGGLAIHNTHLCFVHTV 480
 481 PMDQLFRNPHQALLHTANREDECEVSGGLACHOLCARHGWGPGPTQVNCQFRLGQEC 540
 481 PMDQLFRNPHQALLHTANREDECEVSGGLACHOLCARHGWGPGPTQVNCQFRLGQEC 540
 541 VECRVLQGLPREYVNAHRLPCHPRECPONGSVTCGPEADQCVAAHKDPFCVARC 600
 541 VECRVLQGLPREYVNAHRLPCHPRECPONGSVTCGPEADQCVAAHKDPFCVARC 600
 601 PSGVKDLSMPIMKPFDEBGAQCPICINCHSVDDDDGCGRAEGASLSTISVAVG 660
 601 PSGVKDLSMPIMKPFDEBGAQCPICINCHSVDDDDGCGRAEGASLSTISVAVG 660
 661 ILLVVLGVVFGILLIKRQKIRKTYRRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
 661 ILLVVLGVVFGILLIKRQKIRKTYRRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
 721 RKRVLGSGAGVYVYKIMT PDGENYKIPVALIKYLRNTSPKANKELIDRAYVWAGVSP 780
 721 RKRVLGSGAGVYVYKIMT PDGENYKIPVALIKYLRNTSPKANKELIDRAYVWAGVSP 780
 781 YVERLLGICITSTVQVLTQMLPYGCLLDHYENNGRLSGODLNMCMQIAKMSYLEDVR 840
 781 YVERLLGICITSTVQVLTQMLPYGCLLDHYENNGRLSGODLNMCMQIAKMSYLEDVR 840

841 LVRDLAANVLYKSPNHVKTIDFGARLDIDETRYHADGGKVPKMMALBSILRRFT 900
 841 LVRDLAANVLYKSPNHVKTIDFGARLDIDETRYHADGGKVPKMMALBSILRRFT 900
 901 HQSDVMSYGVWMEIMTEFGAKPYDGI PAREIDLEKGRLLPOPPICTIDVYIMVYKCM 960
 901 HQSDVMSYGVWMEIMTEFGAKPYDGI PAREIDLEKGRLLPOPPICTIDVYIMVYKCM 960
 961 IDSECRPRERELVSERSMARDPQRFVYQNBDELGPASPLDSTFYSLLEDMDGDLVDA 1020
 961 IDSECRPRERELVSERSMARDPQRFVYQNBDELGPASPLDSTFYSLLEDMDGDLVDA 1020
 1021 EHYLVPOGPFCCPDPAAGAGVHHRRSSSTRSGGDLTLGLEPSEBEAPRPLAPSEG 1080
 1021 EHYLVPOGPFCCPDPAAGAGVHHRRSSSTRSGGDLTLGLEPSEBEAPRPLAPSEG 1080
 1081 AGSDVDFGDLGMAKAGLQSLPTHPSPLORYSEDPVLPSETDGVABLTCSPOREYV 1140
 1081 AGSDVDFGDLGMAKAGLQSLPTHPSPLORYSEDPVLPSETDGVABLTCSPOREYV 1140
 1141 NQPDVAPQPPSPREGLPARPAAGATLERPKTISPGKGVYKDVAFAGAVENPEYLTPO 1200
 1141 NQPDVAPQPPSPREGLPARPAAGATLERPKTISPGKGVYKDVAFAGAVENPEYLTPO 1200
 1201 GGAAPQHPPEPAFSPAFLNLYYWDOPPRBGAAPSTFGKPTAEENPEYLGADVPV 1255
 1201 GGAAPQHPPEPAFSPAFLNLYYWDOPPRBGAAPSTFGKPTAEENPEYLGADVPV 1255

RESULT 9
 US-08-484-438-8
 Sequence 8, Application US/08484438
 Patent No. 5811098
 Patent No. 5811098 5780031
 GENERAL INFORMATION:
 APPLICANT: Plowman, Gregory D.
 APPLICANT: Culoscou, Jean-Michel
 APPLICANT: Shoyab, Mohammed
 APPLICANT: Stegall, Clay B.
 APPLICANT: Heller m, Ingegerd
 APPLICANT: Heller m, Karl E.
 TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,438
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/323,442
 FILING DATE: 14-OCT-1994
 APPLICATION NUMBER: US 08/150,704
 FILING DATE: 10-NOV-1993
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/981,165
 FILING DATE: 24-NOV-1992
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Mferock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 5624-230

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIZ
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1255 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-484-438-8

Query Match 98.8%; Score 6732; DB 2; Length 1255;
 Best Local Similarity 99.3%; Pident. No. 0;
 Matches 1247; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

1 MELAALCRWGLLALALPPGAASVQVCTGDMKRLRLPASPETHLDMRLHYOGGVQVQNTL 60
 1 MELAALCRWGLLALALPPGAASVQVCTGDMKRLRLPASPETHLDMRLHYOGGVQVQNTL 60
 61 ELTYLPTNALSFLQDIQEVQGVLLAHNOVROVPLQRLIYVGTQLFEDNYALAVLDNG 120
 61 ELTYLPTNALSFLQDIQEVQGVLLAHNOVROVPLQRLIYVGTQLFEDNYALAVLDNG 120
 121 PLANTPTPTGASPGGLRELQRLSLTEILKGVLLIQRPOLCYQDTILWKDIFHKNOQLA 180
 121 PLANTPTPTGASPGGLRELQRLSLTEILKGVLLIQRPOLCYQDTILWKDIFHKNOQLA 180
 121 PLANTPTPTGASPGGLRELQRLSLTEILKGVLLIQRPOLCYQDTILWKDIFHKNOQLA 180
 181 LTLIDTNSRACHPCSPCKGSRCKGSESDQSLRTVTCAGGCACCKPLPDDCHEQC 240
 181 LTLIDTNSRACHPCSPCKGSRCKGSESDQSLRTVTCAGGCACCKPLPDDCHEQC 240
 241 AAGCTGPKHSIDCLAFHNSGICELHCPALVTYNTDTESPNPEGRYTPGASCTYACP 300
 241 AAGCTGPKHSIDCLAFHNSGICELHCPALVTYNTDTESPNPEGRYTPGASCTYACP 300
 301 YNYLSTDVGSCTLVCPRLHNOEVTABEDGTORCEKSPKPCVCGYGLMEHLREVRATSAN 360
 301 YNYLSTDVGSCTLVCPRLHNOEVTABEDGTORCEKSPKPCVCGYGLMEHLREVRATSAN 360
 301 YNYLSTDVGSCTLVCPRLHNOEVTABEDGTORCEKSPKPCVCGYGLMEHLREVRATSAN 360
 361 IOBFAGCKKIFESLAFLEPSPGDDPASNTAPLOPQLOVFEETLEBTGLYISAMPDSLIP 420
 361 IOBFAGCKKIFESLAFLEPSPGDDPASNTAPLOPQLOVFEETLEBTGLYISAMPDSLIP 420
 361 IOBFAGCKKIFESLAFLEPSPGDDPASNTAPLOPQLOVFEETLEBTGLYISAMPDSLIP 420
 421 DLSPVQNLQVIRGRILHNGAYSLTLQGLISWLGRLSRLRELSGLALIHNNHLLCPVHTV 480
 421 DLSPVQNLQVIRGRILHNGAYSLTLQGLISWLGRLSRLRELSGLALIHNNHLLCPVHTV 480
 421 DLSPVQNLQVIRGRILHNGAYSLTLQGLISWLGRLSRLRELSGLALIHNNHLLCPVHTV 480
 481 PWDOLFNNPHOALLHTANRPEDBCVGEGLACHOLCARHCHWGPPTOCVNCOSQFLRGQEC 540
 481 PWDOLFNNPHOALLHTANRPEDBCVGEGLACHOLCARHCHWGPPTOCVNCOSQFLRGQEC 540
 481 PWDOLFNNPHOALLHTANRPEDBCVGEGLACHOLCARHCHWGPPTOCVNCOSQFLRGQEC 540
 541 VEBGRVIOGLREYVNAHCLPCHPECOPONSVTCFGEADQCYACAHYKOPPCVARC 600
 541 VEBGRVIOGLREYVNAHCLPCHPECOPONSVTCFGEADQCYACAHYKOPPCVARC 600
 541 VEBGRVIOGLREYVNAHCLPCHPECOPONSVTCFGEADQCYACAHYKOPPCVARC 600
 601 PSQVAPDLSYMPIMKFPDEEGACOPCPINCHSCVDLDKGPAPORASPLTSISAVVG 660
 601 PSQVAPDLSYMPIMKFPDEEGACOPCPINCHSCVDLDKGPAPORASPLTSISAVVG 660
 601 PSQVAPDLSYMPIMKFPDEEGACOPCPINCHSCVDLDKGPAPORASPLTSISAVVG 660
 661 ILLVVVLGVVEGIIIRKROQKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILKETEL 720
 661 ILLVVVLGVVEGIIIRKROQKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILKETEL 720
 661 ILLVVVLGVVEGIIIRKROQKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILKETEL 720
 721 RKVVVLGSGAGCTYKGIYIPDGENVKIPVAIKVIRENTSPKANKIIDEAYVMAVGSP 780
 721 RKVVVLGSGAGCTYKGIYIPDGENVKIPVAIKVIRENTSPKANKIIDEAYVMAVGSP 780
 721 RKVVVLGSGAGCTYKGIYIPDGENVKIPVAIKVIRENTSPKANKIIDEAYVMAVGSP 780
 781 YVSHLLGICLSTVQLVTLMPYGLDHRVNRNGRLSODLLMCMQIAGMSTLEVR 840
 781 YVSHLLGICLSTVQLVTLMPYGLDHRVNRNGRLSODLLMCMQIAGMSTLEVR 840
 781 YVSHLLGICLSTVQLVTLMPYGLDHRVNRNGRLSODLLMCMQIAGMSTLEVR 840

QY 841 LVHEDLAARNVLYKSPHVKITDGLARLLDIDETBYHADGKVPKIMMALESILRRRFT 900
 DB 841 LVHEDLAARNVLYKSPHVKITDGLARLLDIDETBYHADGKVPKIMMALESILRRRFT 900
 QY 901 HQSDWMSYGVTTWELMTFGAKPDGIPARIPDLLEGEERLPDPICITDVMYIMVKCM 960
 DB 901 HQSDWMSYGVTTWELMTFGAKPDGIPARIPDLLEGEERLPDPICITDVMYIMVKCM 960
 QY 961 IDSECRPRRELVSERSRMARDPORFVIONEDLGPASPLDSTFYRSLLEDMDGDLVDA 1020
 DB 961 IDSECRPRRELVSERSRMARDPORFVIONEDLGPASPLDSTFYRSLLEDMDGDLVDA 1020
 QY 1021 EBYLVQOQGFPCDPAPGAGVHHRHRSSTSGGDLTLGLEPSEBEAPRSLAPSEG 1080
 DB 1021 EBYLVQOQGFPCDPAPGAGVHHRHRSSTSGGDLTLGLEPSEBEAPRSLAPSEG 1080
 QY 1081 AGSDVDSDGLMGCAAGLOSLPTHDPSPLORYSEDPVPLPSSETDGYVAPLTCSPQPEYV 1140
 DB 1081 AGSDVDSDGLMGCAAGLOSLPTHDPSPLORYSEDPVPLPSSETDGYVAPLTCSPQPEYV 1140
 QY 1141 NOPDVAPOPSPREGPLPARPAGATLERPKTLSPKNGVVKDVFARAGAVENPEYLTPQ 1200
 DB 1141 NOPDVAPOPSPREGPLPARPAGATLERPKTLSPKNGVVKDVFARAGAVENPEYLTPQ 1200
 QY 1201 GGAAPQHPPPAFSPAFDNLYYWDODPPRGAAPPSTFKGPT-AENPEYGLDVPV 1255
 DB 1201 GGAAPQHPPPAFSPAFDNLYYWDODPPRGAAPPSTFKGPT-AENPEYGLDVPV 1255

RESULT 10
 US-09-146-283-4
 Sequence 4, Application US/09146283
 Patent No. 5976546
 GENERAL INFORMATION:
 APPLICANT: Laus, Reiner
 APPLICANT: Ruegg, Curtis L.
 APPLICANT: Wu, Hongyu
 TITLE OF INVENTION: Immunostimulatory Compositions
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Denlinger & Associates
 STREET: 350 Cambridge Ave. Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/146,283
 FILING DATE: 03-SEPT-1998
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Judge, Linda R.
 REGISTRATION NUMBER: 42,702
 REFERENCE/DOCKET NUMBER: 7636-0010.21
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 782 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: homo sapiens
 INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
 US-09-146-283-4

Query Match 53.3%; Score 3630; DB 2; Length 782;
 Best Local Similarity 99.8%; Pred. No. 1.2e-238;
 Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELAALCRWGLLALLPPGAASVQVCTGTDKRLPASPTHTLMLHLVYOGCOVQGNL 60
DB 1 MELAALCRWGLLALLPPGAASVQVCTGTDKRLPASPTHTLMLHLVYOGCOVQGNL 60
QY 61 ELYTLPTNASLSFLQDIQEVQGYVLIANQVQVPLQRLAIYVGTQVLFEDNYALAVDNG 120
DB 61 ELYTLPTNASLSFLQDIQEVQGYVLIANQVQVPLQRLAIYVGTQVLFEDNYALAVDNG 120
QY 121 DPLNNTTPTVGSFGLRELQRLSTELKGVLIQNPOLCYQDTILMKDIFKNNOLA 180
DB 121 DPLNNTTPTVGSFGLRELQRLSTELKGVLIQNPOLCYQDTILMKDIFKNNOLA 180
QY 181 LTLIDNRSBACHPCSPCKGSRGWSSESDCQSLTRTVACAGGACRCKGPLPTDCHEQC 240
DB 181 LTLIDNRSBACHPCSPCKGSRGWSSESDCQSLTRTVACAGGACRCKGPLPTDCHEQC 240
QY 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYACP 300
DB 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYACP 300
QY 301 YNLTSTDVGSCTIVCPILHNOEVTABDGTQRCCKSKFCARVCYGLGMEHLREVAVTSAN 360
DB 301 YNLTSTDVGSCTIVCPILHNOEVTABDGTQRCCKSKFCARVCYGLGMEHLREVAVTSAN 360
QY 361 IQEPAGCKKI FGSIAFLPESPDGPASNTAPLOPEQLOVEFTELEITGYLIYSMPSLP 420
DB 361 IQEPAGCKKI FGSIAFLPESPDGPASNTAPLOPEQLOVEFTELEITGYLIYSMPSLP 420
QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRSRLRELSGLALIHNTHLCPVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRSRLRELSGLALIHNTHLCPVHTV 480
QY 481 PMDLFNPFOALHTANRPEDCEVGEGLACHOLCARGHCWGPPTCVNCSOFLRQEC 540
DB 481 PMDLFNPFOALHTANRPEDCEVGEGLACHOLCARGHCWGPPTCVNCSOFLRQEC 540
QY 541 VESCRVLOGLPREVYNAHCLPCHPECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600
DB 541 VESCRVLOGLPREVYNAHCLPCHPECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600
QY 601 PGGVKPDLSTMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCPRAEQASPLTSL 654
DB 601 PGGVKPDLSTMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCPRAEQASPLTSL 654

```

RESULT 11

US-08-579-823A-4
 ; Sequence 4, Application US/08579823A
 ; Patent No. 6080409
 ; GENERAL INFORMATION:
 ; APPLICANT: Laue, Reiner
 ; APPLICANT: Ruegg, Curie L.
 ; APPLICANT: Wu, Hongyu
 ; TITLE OF INVENTION: Immunostimulatory Composition and Method
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Delinger & Associates
 ; STREET: 350 Cambridge Ave. Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/579, 823A
 FILING DATE: 03-DEC-1998
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Judge, Linda R.
 REGISTRATION NUMBER: 42,702
 REFERENCE/DOCKET NUMBER: 7636-0010
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 782 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: homo sapiens
 INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
 US-08-579-823A-4

Query Match 53.3%; Score 3630; DB 3; Length 782;
 Best Local Similarity 99.8%; Pred. No. 1.2e-238;
 Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELAALCRWGLLALLPPGAASVQVCTGTDKRLPASPTHTLMLHLVYOGCOVQGNL 60
DB 1 MELAALCRWGLLALLPPGAASVQVCTGTDKRLPASPTHTLMLHLVYOGCOVQGNL 60
QY 61 ELYTLPTNASLSFLQDIQEVQGYVLIANQVQVPLQRLAIYVGTQVLFEDNYALAVDNG 120
DB 61 ELYTLPTNASLSFLQDIQEVQGYVLIANQVQVPLQRLAIYVGTQVLFEDNYALAVDNG 120
QY 121 DPLNNTTPTVGSFGLRELQRLSTELKGVLIQNPOLCYQDTILMKDIFKNNOLA 180
DB 121 DPLNNTTPTVGSFGLRELQRLSTELKGVLIQNPOLCYQDTILMKDIFKNNOLA 180
QY 181 LTLIDNRSBACHPCSPCKGSRGWSSESDCQSLTRTVACAGGACRCKGPLPTDCHEQC 240
DB 181 LTLIDNRSBACHPCSPCKGSRGWSSESDCQSLTRTVACAGGACRCKGPLPTDCHEQC 240
QY 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYACP 300
DB 241 AACCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTPESMPREGRYTFGASCTYACP 300
QY 301 YNLTSTDVGSCTIVCPILHNOEVTABDGTQRCCKSKFCARVCYGLGMEHLREVAVTSAN 360
DB 301 YNLTSTDVGSCTIVCPILHNOEVTABDGTQRCCKSKFCARVCYGLGMEHLREVAVTSAN 360
QY 361 IQEPAGCKKI FGSIAFLPESPDGPASNTAPLOPEQLOVEFTELEITGYLIYSMPSLP 420
DB 361 IQEPAGCKKI FGSIAFLPESPDGPASNTAPLOPEQLOVEFTELEITGYLIYSMPSLP 420
QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRSRLRELSGLALIHNTHLCPVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRSRLRELSGLALIHNTHLCPVHTV 480
QY 481 PMDLFNPFOALHTANRPEDCEVGEGLACHOLCARGHCWGPPTCVNCSOFLRQEC 540
DB 481 PMDLFNPFOALHTANRPEDCEVGEGLACHOLCARGHCWGPPTCVNCSOFLRQEC 540
QY 541 VESCRVLOGLPREVYNAHCLPCHPECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600
DB 541 VESCRVLOGLPREVYNAHCLPCHPECOPONGSVTCGPEADQCVACAHYKDPFCVARC 600
QY 601 PGGVKPDLSTMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCPRAEQASPLTSL 654
DB 601 PGGVKPDLSTMPIMKFPDEEGACQPCPINCTHSCVDLDDKGCPRAEQASPLTSL 654

```

RESULT 12

US-09-344-195-4

Sequence 4, Application US/09344195
Patent No. 6210662
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
Mu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,195
FILING DATE: 24-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-344-195-4
Query Match 53.3%; Score 3630; DB 3; Length 782;
Best Local Similarity 99.8%; Pred. No. 1,2e-238;
Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLALPPGAASVQCTGDMKRLPASPEHLDMLRHLYOGQVQVQNL 60
DB 1 MELAALCRWGLLALPPGAASVQCTGDMKRLPASPEHLDMLRHLYOGQVQVQNL 60
QY 61 ELTYIPTNASLFLDIOEVQGVILAHNOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
DB 61 ELTYIPTNASLFLDIOEVQGVILAHNOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTGASPGGLRELQRLSLREILKGVILIRNQLCYDTITLWKDIFHKNNOLA 180
DB 121 DPLNNTTPTGASPGGLRELQRLSLREILKGVILIRNQLCYDTITLWKDIFHKNNOLA 180
QY 181 LTLIDTNSRACHPCSPCKGSRGWSGSSBDOSLTRVCAAGCARCKGFLPTDCCHQC 240
DB 181 LTLIDTNSRACHPCSPCKGSRGWSGSSBDOSLTRVCAAGCARCKGFLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACHFHNSGICELHCPALVYNTDIFESMPNPEGRTYFGASCTYAC 300
DB 241 AAGCTGPKHSDCLACHFHNSGICELHCPALVYNTDIFESMPNPEGRTYFGASCTYAC 300
QY 301 YNYLSTDVGSCTLVCPHNOEVTAEQGTORCEKSKPCARVCYGLGNEHLREAVATYSAN 360
DB 301 YNYLSTDVGSCTLVCPHNOEVTAEQGTORCEKSKPCARVCYGLGNEHLREAVATYSAN 360

DB 301 YNYLSTDVGSCTLVCPHNOEVTAEQGTORCEKSKPCARVCYGLGNEHLREAVATYSAN 360
QY 361 IOEFAGCKKIFGSLAFLPESFPDGPASNTAPLOPEQLOVEFTELETGYLISAMPDLP 420
DB 361 IOEFAGCKKIFGSLAFLPESFPDGPASNTAPLOPEQLOVEFTELETGYLISAMPDLP 420
QY 421 DISVFONLQVTRRIILHNGVSLTLOGLISWLGRLSRLGSLALIHNTHLCPVHTV 480
DB 421 DISVFONLQVTRRIILHNGVSLTLOGLISWLGRLSRLGSLALIHNTHLCPVHTV 480
QY 481 PMQLFRNPQALHTANRPEDSCVGBGLAQLCARGHGWSGPGPTCCVNCSTLRGQEC 540
DB 481 PMQLFRNPQALHTANRPEDSCVGBGLAQLCARGHGWSGPGPTCCVNCSTLRGQEC 540
QY 541 VEECRVLOGLPREYVNAARHCLPCHPECQPNQSGSVTCGFPADQVACAHYKDPFCVARC 600
DB 541 VEECRVLOGLPREYVNAARHCLPCHPECQPNQSGSVTCGFPADQVACAHYKDPFCVARC 600
QY 601 PSCVKEPDLSTMPIMKFPDEBGACQPCPINCTHSCVDLDDKGCFAEORASPLTSI 654
DB 601 PSCVKEPDLSTMPIMKFPDEBGACQPCPINCTHSCVDLDDKGCFAEORASPLTSI 654
RESULT 13
US-08-422-108-1
Sequence 1, Application US/08422108
Patent No. 6015567
GENERAL INFORMATION:
APPLICANT: Hudziak, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,108
FILING DATE: 14-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/355460
FILING DATE: 13-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/354319
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C2D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-08-422-108-1

Query Match 51.0%; Score 3473; DB 3; Length 624;
 Best Local Similarity 99.8%; Pred. No. 4.2e-228;
 Matches 623; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 22 STQVCTGTDMKRLPASPETHLMDLRLHYGCGVVOGNIETLYLPTNASLSPLODIOEVQ 81
DB 1 STQVCTGTDMKRLPASPETHLMDLRLHYGCGVVOGNIETLYLPTNASLSPLODIOEVQ 60
QY 82 GYVLIANOVROVPLQRLRIVRGTOIFEDNYALAVLDNGPLNNTPVTGASPGGLRELO 141
DB 61 GYVLIANOVROVPLQRLRIVRGTOIFEDNYALAVLDNGPLNNTPVTGASPGGLRELO 120
QY 142 LNSLTELKGGVLIQIRNPOLCYODTILMKDIFKNNQALATLIDTNSRACHPCSPCKG 201
DB 121 LNSLTELKGGVLIQIRNPOLCYODTILMKDIFKNNQALATLIDTNSRACHPCSPCKG 180
QY 202 SRMGSSSEDCQSLRTVCAAGGACRCKGPLPTDCHEQCAAGCTGPRHSDCLCLHFNHS 261
DB 181 SRMGSSSEDCQSLRTVCAAGGACRCKGPLPTDCHEQCAAGCTGPRHSDCLCLHFNHS 240
QY 262 GICELHCPALVTYNTDTFESMPNBEGRYTFGASCVTACPNYISTDVGSCTLVCPPLHNOE 321
DB 241 GICELHCPALVTYNTDTFESMPNBEGRYTFGASCVTACPNYISTDVGSCTLVCPPLHNOE 300
QY 322 VTAEDEGTORCEKSKPCARVCYGLGMEHLREVAVTSANTIOEPAGCKKIFGSLAFLEPSF 381
DB 301 VTAEDEGTORCEKSKPCARVCYGLGMEHLREVAVTSANTIOEPAGCKKIFGSLAFLEPSF 360
QY 382 DQDPASNTAPLOEPOLOVFETLEITGYLIASMPDLPDLSVQNLQVIRGRILHNGAY 441
DB 361 DQDPASNTAPLOEPOLOVFETLEITGYLIASMPDLPDLSVQNLQVIRGRILHNGAY 420
QY 442 SLTLQGLISWLGRLSRLREISGLALIHNTHLCFVHTVPMDOLEFRPHQALHTANRPE 501
DB 421 SLTLQGLISWLGRLSRLREISGLALIHNTHLCFVHTVPMDOLEFRPHQALHTANRPE 480
QY 502 DRCVGBGLACHOLCARGHCWPGPTQCVNCSOPFRGQECYBECKVLOGLPREYVNAHCL 561
DB 481 DRCVGBGLACHOLCARGHCWPGPTQCVNCSOPFRGQECYBECKVLOGLPREYVNAHCL 540
QY 562 PCHPECOPOKGSVTCFGEPAADQCAAHYKDPPECVACRCSGVKPDLSYPMKFPDEEG 621
DB 541 PCHPECOPOKGSVTCFGEPAADQCAAHYKDPPECVACRCSGVKPDLSYPMKFPDEEG 600
QY 622 ACQPCPIINCTHSCVLDLDDKCPAE 645
DB 601 ACQPCPIINCTHSCVLDLDDKCPAE 624

```

RESULT 14
 US-08-422-734-1
 Sequence 1, Application US/08422734
 Patent No. 633169
 GENERAL INFORMATION:
 APPLICANT: Huddiak, Robert M.
 APPLICANT: Shepard, H. Michael
 APPLICANT: Ullrich, Axel
 TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/422,734

FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/422108
 FILING DATE: 14-Apr-1995
 APPLICATION NUMBER: 08/355460
 FILING DATE: 13-DEC-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/048346
 FILING DATE: 15-APR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/354319
 FILING DATE: 19-MAY-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Lee, Wendy M
 REGISTRATION NUMBER: 00,000
 REFERENCE/DOCKET NUMBER: 554C201
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-1994
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 624 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear

Query Match 51.0%; Score 3473; DB 4; Length 624;
 Best Local Similarity 99.8%; Pred. No. 4.2e-228;
 Matches 623; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 22 STQVCTGTDMKRLPASPETHLMDLRLHYGCGVVOGNIETLYLPTNASLSPLODIOEVQ 81
DB 1 STQVCTGTDMKRLPASPETHLMDLRLHYGCGVVOGNIETLYLPTNASLSPLODIOEVQ 60
QY 82 GYVLIANOVROVPLQRLRIVRGTOIFEDNYALAVLDNGPLNNTPVTGASPGGLRELO 141
DB 61 GYVLIANOVROVPLQRLRIVRGTOIFEDNYALAVLDNGPLNNTPVTGASPGGLRELO 120
QY 142 LNSLTELKGGVLIQIRNPOLCYODTILMKDIFKNNQALATLIDTNSRACHPCSPCKG 201
DB 121 LNSLTELKGGVLIQIRNPOLCYODTILMKDIFKNNQALATLIDTNSRACHPCSPCKG 180
QY 202 SRMGSSSEDCQSLRTVCAAGGACRCKGPLPTDCHEQCAAGCTGPRHSDCLCLHFNHS 261
DB 181 SRMGSSSEDCQSLRTVCAAGGACRCKGPLPTDCHEQCAAGCTGPRHSDCLCLHFNHS 240
QY 262 GICELHCPALVTYNTDTFESMPNBEGRYTFGASCVTACPNYISTDVGSCTLVCPPLHNOE 321
DB 241 GICELHCPALVTYNTDTFESMPNBEGRYTFGASCVTACPNYISTDVGSCTLVCPPLHNOE 300
QY 322 VTAEDEGTORCEKSKPCARVCYGLGMEHLREVAVTSANTIOEPAGCKKIFGSLAFLEPSF 381
DB 301 VTAEDEGTORCEKSKPCARVCYGLGMEHLREVAVTSANTIOEPAGCKKIFGSLAFLEPSF 360
QY 382 DQDPASNTAPLOEPOLOVFETLEITGYLIASMPDLPDLSVQNLQVIRGRILHNGAY 441
DB 361 DQDPASNTAPLOEPOLOVFETLEITGYLIASMPDLPDLSVQNLQVIRGRILHNGAY 420
QY 442 SLTLQGLISWLGRLSRLREISGLALIHNTHLCFVHTVPMDOLEFRPHQALHTANRPE 501
DB 421 SLTLQGLISWLGRLSRLREISGLALIHNTHLCFVHTVPMDOLEFRPHQALHTANRPE 480
QY 502 DRCVGBGLACHOLCARGHCWPGPTQCVNCSOPFRGQECYBECKVLOGLPREYVNAHCL 561
DB 481 DRCVGBGLACHOLCARGHCWPGPTQCVNCSOPFRGQECYBECKVLOGLPREYVNAHCL 540
QY 562 PCHPECOPOKGSVTCFGEPAADQCAAHYKDPPECVACRCSGVKPDLSYPMKFPDEEG 621
DB 541 PCHPECOPOKGSVTCFGEPAADQCAAHYKDPPECVACRCSGVKPDLSYPMKFPDEEG 600
QY 622 ACQPCPIINCTHSCVLDLDDKCPAE 645

```

Db 601 ACOPPCINCHSCVDLDDKCCPAB 624

RESULT 15

US-08-475-035-4
Sequence 4, Application US/08475035
Patent No. 5985553

GENERAL INFORMATION:

APPLICANT: KING, C. R.
APPLICANT: KRAUS, MATTHIAS H.
APPLICANT: AARONSON, STUART A.
TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
TITLE OF INVENTION: EGF RECEPTOR GENE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: NEEDLE & ROSENBERG, P.C.
STREET: Suite 1200, 127 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475, 035
FILING DATE: 7 Jun 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414.656
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1210 amino acids
TYPE: amino acid
TOPOLOGY: 1linear
MOLECULE TYPE: protein

US-08-475-035-4

Query Match 46.5%; Score 3169; DB 2; Length 1210;
Best Local Similarity 49.8%; Pred. No. 4, 5e-207;
Matches 632; Conservative 177; Mismatches 345; Indels 116; Gaps 22;

QY 11 LLLALLPPGAA--STOVCTGTDMKRLTPASPTHLDMLEHLVYOGCOVGNLELTPTN 68
DB 14 LLAALCPASBALBEKKVCGSTNKLTOIGTFEDHLSLORMFNCEVVLGNLEITTVQNR 73
QY 69 ASLSLDOIIOYQGYVLAHNOVROPLOKRLIVRGTOLEFEDNALAVLNDGDPINNTTP 128
DB 74 YDLSFLKIOBAGVYVLAALNVERIIPLENQIIRGNMYENSVALAVLSND----- 126
QY 129 VTGASPGSLRELOLSLTELKGVLIQNPOLCYQDTILMDIETHKNNLALTLIDNR 188
DB 127 ---ANKTGLKEIPMNLOEILHGAVRFSNNPALCNVESTIQMRDIVSPFLSNMSDFQNR 183
QY 189 SRACHPCSPMGKSKWGESSEDCOSLTRTVACAGCA--RCKGPLPTDCCBECAGCTGP 247
DB 184 LGSQCKDPSCPNCSCKWAGEENCKLTKIICAQCCSCKRCKSKSDCCCHNCAAGCTGP 243
QY 248 KHSQCLALHNRHSGICELHCPALVTYNTDFESMNPDEGRATFGASCCTACPYNYLSTD 307
DB 244 RESDCLVCRKRFRDEATKCTCPPLMLYNPTTYQMDVNPEGKYSFGATCVKCKPRYVYVTD 303
QY 308 VGSCTLVPLINQETATADGTCORCEKSKPCARVYCYGLGHEHLREVRATVANSNIOEPAGC 367
DB 304 HGSCTVACGADSYEM--BEDGVAKCKCKGCPCKRVKCNIGIGIFKDSLSINNTINIGIFKNC 362

QY 368 KTFGSLAFLPSFDDDPASNTAPLOBOLOVEFTELEITGYLYISAMPDPLPSYFON 427
DB 363 TSIISDHLILPVAFRDSDFTHTPPLDPOBELDIKTYKEITGFLILOMFERITDLHAFEN 422
QY 428 LOYIRGRILHNGAVSYLTLOGLSISWLSLRELSGLALIHNTLCPVHTVMDOLF 487
DB 423 LEIIRGRKQHQGFSLAVVSLNITSLSLRSLSKESISDGDVLIISGNKNCVANTIMWKLFG 482
QY 488 NPHQALHTANPEDECSVEGLACHOLCARHGWCGPPOCVNCSQSLROQCEBERVL 547
DB 483 TSGQTKIISNNGENSCAKTGOVCHALCPBEGWGEPRDVCSCRVNRRRECDKCKL 542
QY 548 OGCPREYVNAHCLPCHPECOPONGSVTCFGEPAQCVACAHKDPFCVAPRCSPGKPD 607
DB 543 EGSPREFEVNSCICQHPBCLDQAMNITCTGAPNDICQCAHIIDGPHCVKCTPAGWGE 602
QY 608 LSTMPWKFPDEBAGCCPINCCHSCVDLDDKCPAQRASPLTJISAVG--LLV 664
DB 603 NNLT--VMKYADAGHVCHLCPNCTYCTGPGLEGCPNGPKIP--SIATGMVGLLLLV 659
QY 665 VTLGVVFGILIRROOKIRKTYMRRLQETELVEPLTBSGAMNOQMRITKTELKVK 724
DB 660 VALGIG--LEMRRRRIYAKRTLRLLORELEVEPLTBSGABNOALRLTKEFEKXIK 716
QY 725 VLSGAFGTVYKGIWIPDGENYKIPVAIKVLRBNTSPKANKETLDEAYVWAGSPVSR 784
DB 717 VLSGAFGTVYKGIWIPBEEKKIPVAIKELBATSPPANKETLDEAYVWASVDNPHVCR 776
QY 785 LIGICLSTVQLVTOAMPYGLLDHYRENRGLSGODLNMCMQIAKNGSYLDEVLVHR 844
DB 777 ILGICLSTVQLVTOAMPYGLLDHYRENRGLSGODLNMCMQIAKNGSYLDEVLVHR 836
QY 845 DLAARNVLYKSPNHYKITDPGLARLIDETVHADGCVPIKMALESILRRRPTHQSD 904
DB 837 DLAARNVLYKTPQHYKITDPGLAKLIGABEKHYHGGKVPKIMALESILHRIYTHQSD 896
QY 905 VMSYGVYVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYIMVYKMMIDSE 964
DB 897 VMSYGVYVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYIMVYKMMIDSE 956
QY 965 CRPRELIVSEFSRNARDPQRFVVIQ--NBDLQPASPLDSTFYRSLDEDDMDGLVDAEY 1023
DB 957 SRPKFRELILERSKVARDPQRYLVIQSDBRMHLPSPTDSNFYRLMDEEDMDVVADEX 1016
QY 1024 LVPQGFCCPDPAFGAGMHHRRHSSSTRSGGDLTLGLEPSEBEAPSPPLAPSGAGS 1083
DB 1017 LIPQGF--SSPSTSRTPILLSLSATS 1042
QY 1084 DVFDGDLGMAKGLQSLPTHDPSLOYSDPTVPLPSFT--DGVAVPLTCSPOPEYVN 1141
DB 1043 N--NSTVACIDRNGIQSCPIYEDSFLQRYSDPTALTEDSIDDFL-----PVEYIN 1094
QY 1142 QPDVAPPOPSPREGLPAPAPAGATL-----ERPXTSPGKGVVVDVAFGAGAVENPEY 1196
DB 1095 Q-----SVF--KRPAGSYQNPVYHNPQPLNAPASRPDPHOD--PHSTAVGNPEY 1138
QY 1197 L--TPOGGAAPPHPPAPSPAPNDLYWDQ-----DP-----PERGAPSTFGK 1239
DB 1139 LNTVO-----FTCVNSTFDSFAPMAOKSHQISIDNPYQODFFPKAKPNCIGIFK 1189
QY 1240 TPTAENPEYL 1249
DB 1190 S-TAENPEYL 1198

Search completed: February 9, 2004, 16:17:57
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 7, 2004, 13:30:29 ; Search time 34 Seconds

(Without alignments)
61.583 Million cell updates/sec

Title: US-09-930-125-3

Perfect score: 54

Sequence: 1 EBYLVPQGGF 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA.*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	10	US-09-930-125-3	Sequence 3, Appl1
2	54	100.0	10	US-10-313-644-3	Sequence 3, Appl1
3	54	100.0	59	US-09-854-356-5	Sequence 5, Appl1
4	54	100.0	266	US-09-854-356-7	Sequence 4, Appl1
5	54	100.0	293	US-10-102-806-583	Sequence 583, Appl1
6	54	100.0	570	US-10-378-393-15	Sequence 15, Appl1
7	54	100.0	583	US-09-930-125-9	Sequence 9, Appl1
8	54	100.0	589	US-09-930-125-8	Sequence 8, Appl1
9	54	100.0	589	US-09-930-125-10	Sequence 10, Appl1
10	54	100.0	600	US-09-930-125-11	Sequence 11, Appl1
11	54	100.0	712	US-09-854-356-7	Sequence 7, Appl1
12	54	100.0	791	US-10-378-393-11	Sequence 11, Appl1
13	54	100.0	919	US-09-854-356-6	Sequence 6, Appl1
14	54	100.0	1253	US-10-146-473-72	Sequence 72, Appl1
15	54	100.0	1255	US-09-811-123-9	Sequence 9, Appl1

16	54	100.0	1255	9	US-09-811-115-3	Sequence 3, Appl1
17	54	100.0	1255	10	US-09-769-508-2	Sequence 2, Appl1
18	54	100.0	1255	10	US-09-854-356-1	Sequence 1, Appl1
19	54	100.0	1255	10	US-09-930-125-2	Sequence 2, Appl1
20	54	100.0	1255	11	US-09-441-411-6	Sequence 6, Appl1
21	54	100.0	1255	12	US-10-207-498-6	Sequence 6, Appl1
22	54	100.0	1255	12	US-10-338-730-2	Sequence 2, Appl1
23	54	100.0	1255	12	US-10-313-644-2	Sequence 2, Appl1
24	54	100.0	1255	12	US-10-322-892-4	Sequence 2, Appl1
25	54	100.0	1255	12	US-10-272-437A-28	Sequence 28, Appl1
26	54	100.0	1255	12	US-10-117-937-59A	Sequence 59A, Appl1
27	54	100.0	1255	12	US-10-394-322A-17	Sequence 17, Appl1
28	54	100.0	1255	12	US-10-245-871-553	Sequence 553, Appl1
29	54	100.0	1255	12	US-10-149-138-4641	Sequence 4641, Appl1
30	54	100.0	1255	12	US-10-435-696-36	Sequence 36, Appl1
31	54	100.0	1255	15	US-10-207-655-45	Sequence 45, Appl1
32	54	100.0	1255	15	US-10-177-293-126	Sequence 126, Appl1
33	54	100.0	1256	10	US-09-854-356-2	Sequence 2, Appl1
34	54	100.0	1256	10	US-09-854-356-14	Sequence 14, Appl1
35	54	100.0	1260	10	US-09-870-759-118	Sequence 118, Appl1
36	54	100.0	1260	12	US-09-751-708A-118	Sequence 118, Appl1
37	50	92.6	20	12	US-10-238-960-5	Sequence 5, Appl1
38	50	92.6	20	12	US-10-238-960-6	Sequence 6, Appl1
39	50	92.6	1210	9	US-09-725-433-2	Sequence 2, Appl1
40	50	92.6	1210	12	US-10-394-322A-16	Sequence 16, Appl1
41	49	90.7	9	12	US-10-245-871-607	Sequence 607, Appl1
42	49	90.7	9	12	US-10-149-138-1138	Sequence 1138, Appl1
43	49	90.7	9	12	US-10-149-138-3666	Sequence 3666, Appl1
44	49	90.7	10	12	US-10-149-138-1139	Sequence 1139, Appl1
45	49	90.7	10	12	US-10-149-138-3667	Sequence 3667, Appl1

ALIGNMENTS

RESULT 1
US-09-930-125-3
; Sequence 3, Application US/09930125
; Publication No. US20020193329A1
GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; TYPE: PRT
; LENGTH: 10
; ORGANISM: Homo sapiens
US-09-930-125-3

Query Match 100.0% Score 54; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00068;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EBYLVPQGGF 10
Db 1 EBYLVPQGGF 10

RESULT 2
US-10-313-644-3
; Sequence 3, Application US/10313644
; Publication No. US20030157119A1

GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Cheever, Martin A.
APPLICANT: Hand-Zimmerman, Susan
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
FILE REFERENCE: 210121.483C3
CURRENT APPLICATION NUMBER: US/10/313,644
CURRENT FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapiens
US-10-313-644-3

Query Match 100.0%; Score 54; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10
DB 1 EBYLVPQGGF 10

RESULT 3
US-09-854-356-5
Sequence 5; Application US/09854356-5
Patent No. US2002017567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheever, Dik
APPLICANT: Corixa Corporation
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
CURRENT FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: US 09/493,480
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,976
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD) of human HER-2/neu
US-09-854-356-5

Query Match 100.0%; Score 54; DB 10; Length 59;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10
DB 32 EBYLVPQGGF 41

RESULT 4
US-09-854-356-4
Sequence 4; Application US/09854356-4
Patent No. US2002017567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheever, Dik
APPLICANT: Corixa Corporation
TITLE OF INVENTION: HER-2/neu Fusion Proteins

FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
CURRENT FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: US 09/493,480
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,976
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 266
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
US-09-854-356-4

Query Match 100.0%; Score 54; DB 10; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPQGGF 10
DB 32 EBYLVPQGGF 41

RESULT 5
US-10-102-806-583
Sequence 583; Application US/10102806
Publication No. US2003005442A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103PIC1
CURRENT APPLICATION NUMBER: US/10/102,806
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/925,298
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 583
LENGTH: 293
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (52)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (53)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (58)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (150)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (171)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (207)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (254)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-583

Query Match 100.0%; Score 54; DB 15; Length 293;

Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 105 BEYLVPOQGF 114

RESULT 6
US-10-378-393-15

/ Sequence 15, Application US/10378393
/ Publication No. US20030182668A1
/ GENERAL INFORMATION:

/ APPLICANT: Bol, David K.
/ APPLICANT: Carboni, Joan M.
/ APPLICANT: Rowley, Ronald B.
/ APPLICANT: Wong, Tai W.
/ APPLICANT: Lee, Francis Y.
/ TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
/ FILE REFERENCE: D0254 NP
/ CURRENT APPLICATION NUMBER: US/10/378,393
/ CURRENT FILING DATE: 2003-03-03
/ PRIOR APPLICATION NUMBER: US 60/360,889
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 15
/ LENGTH: 570
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-378-393-15

Query Match 100.0%; Score 54; DB 12; Length 570;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 336 BEYLVPOQGF 345

RESULT 7
US-09-930-125-9

/ Sequence 9, Application US/09930125
/ Publication No. US20020193329A1
/ GENERAL INFORMATION:

/ APPLICANT: Hand-Zimmerman, Susan
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Foy, Teresa M.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
/ FILE REFERENCE: 210121.544
/ CURRENT APPLICATION NUMBER: US/09/930,125
/ CURRENT FILING DATE: 2001-08-14
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 9
/ LENGTH: 583
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-930-125-9

Query Match 100.0%; Score 54; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 105 BEYLVPOQGF 114

Db 347 BEYLVPOQGF 356

RESULT 8
US-09-930-125-8

/ Sequence 8, Application US/09930125
/ Publication No. US20020193329A1
/ GENERAL INFORMATION:

/ APPLICANT: Hand-Zimmerman, Susan
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Foy, Teresa M.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
/ FILE REFERENCE: 210121.544
/ CURRENT APPLICATION NUMBER: US/09/930,125
/ CURRENT FILING DATE: 2001-08-14
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 8
/ LENGTH: 587
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-930-125-8

Query Match 100.0%; Score 54; DB 10; Length 587;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 347 BEYLVPOQGF 356

RESULT 9
US-09-930-125-10

/ Sequence 10, Application US/09930125
/ Publication No. US20020193329A1
/ GENERAL INFORMATION:

/ APPLICANT: Hand-Zimmerman, Susan
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Foy, Teresa M.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
/ FILE REFERENCE: 210121.544
/ CURRENT APPLICATION NUMBER: US/09/930,125
/ CURRENT FILING DATE: 2001-08-14
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 10
/ LENGTH: 589
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-930-125-10

Query Match 100.0%; Score 54; DB 10; Length 589;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEYLVPOQGF 10
Db 355 BEYLVPOQGF 364

RESULT 10
US-09-930-125-11

```
/ Sequence 11, Application US/09930125
/ Publication No. US20020193329A1
/ GENERAL INFORMATION:
/ APPLICANT: Hand-Zimmerman, Susan
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Roy, Teresa M.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Vedvick, Thomas S.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
/ TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
/ FILE REFERENCE: 210121.544
/ CURRENT APPLICATION NUMBER: US/09/930.125
/ CURRENT FILING DATE: 2001-08-14
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 600
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-930-125-11

Query Match      100.0%; Score 54; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
Db 366 EBYLVPOQGF 375

RESULT 11
US-09-854-356-7
/ Sequence 7, Application US/09854356
/ Patent No. US2002017567A1
/ GENERAL INFORMATION:
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Gheysen, Dirk
/ APPLICANT: Corixa Corporation
/ APPLICANT: SmithKline Beecham Biologicals S. A.
/ TITLE OF INVENTION: HER-2/neu Fusion Proteins
/ FILE REFERENCE: 014058-009810PC
/ CURRENT APPLICATION NUMBER: US/09/854.356
/ CURRENT FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 09/493,480
/ PRIOR FILING DATE: 2000-01-28
/ PRIOR APPLICATION NUMBER: US 60/117,976
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 7
/ LENGTH: 712
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: fusion protein
/ OTHER INFORMATION: of ECD and delta PD of human HER-2/neu
US-09-854-356-7

Query Match      100.0%; Score 54; DB 10; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
Db 685 EBYLVPOQGF 694

RESULT 12
US-10-378-393-11
/ Sequence 11, Application US/10378393
/ Publication No. US2003018268A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Bol, David K.
/ APPLICANT: Cardoni, Joan M.
/ APPLICANT: Rowley, Ronald B.
/ APPLICANT: Wong, Tai W.
/ APPLICANT: Lee, Francis Y.
/ TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
/ TITLE OF INVENTION: TYROSINE KINASE RECEPTORS
/ FILE REFERENCE: D0254 NE
/ CURRENT APPLICATION NUMBER: US/10/378.393
/ CURRENT FILING DATE: 2003-03-03
/ PRIOR APPLICATION NUMBER: US 60/360,889
/ PRIOR FILING DATE: 2002-03-01
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 11
/ LENGTH: 791
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-378-393-11

Query Match      100.0%; Score 54; DB 12; Length 791;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
Db 557 EBYLVPOQGF 566

RESULT 13
US-09-854-356-6
/ Sequence 6, Application US/09854356
/ Patent No. US2002017567A1
/ GENERAL INFORMATION:
/ APPLICANT: Cheever, Martin A.
/ APPLICANT: Gheysen, Dirk
/ APPLICANT: Corixa Corporation
/ APPLICANT: SmithKline Beecham Biologicals S. A.
/ TITLE OF INVENTION: HER-2/neu Fusion Proteins
/ FILE REFERENCE: 014058-009810PC
/ CURRENT APPLICATION NUMBER: US/09/854.356
/ CURRENT FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 09/493,480
/ PRIOR FILING DATE: 2000-01-28
/ PRIOR APPLICATION NUMBER: US 60/117,976
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 919
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: fusion protein
/ OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6

Query Match      100.0%; Score 54; DB 10; Length 919;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EBYLVPOQGF 10
Db 685 EBYLVPOQGF 694

RESULT 14
US-10-146-473-72
/ Sequence 72, Application US/10146473
/ Publication No. US2003010888A1
/ GENERAL INFORMATION:
/ APPLICANT: Scanlan, Matthew
```

```

; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Cure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 1253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-72

```

```

Query Match      100.0%; Score 54; DB 15; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 EBYLVPOQGF 10
Db      1021 EBYLVPOQGF 1030

```

```

RESULT 15
US-09-811-123-9
; Sequence 9, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwkowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EPDB
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-123-9

```

```

Query Match      100.0%; Score 54; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 EBYLVPOQGF 10
Db      1021 EBYLVPOQGF 1030

```

Search completed: February 7, 2004, 13:34:01
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 16:16:50 ; Search time 45 Seconds
(without alignments)
5839.439 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815
Sequence: 1 METALCRWGLLALLPGCA.....TRKGTPTANPEYLGIDVTV 1255

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6815	100.0	1255	10 US-09-854-356-1	Sequence 1, Appl1
2	6815	100.0	1255	10 US-09-930-125-2	Sequence 2, Appl1
3	6815	100.0	1255	11 US-09-441-411-6	Sequence 6, Appl1
4	6815	100.0	1255	11 US-10-313-644-2	Sequence 2, Appl1
5	6815	100.0	1255	12 US-10-394-322A-17	Sequence 17, Appl1
6	6815	100.0	1255	12 US-10-245-871-553	Sequence 33, Appl1
7	6815	100.0	1255	13 US-10-207-655-45	Sequence 45, Appl1
8	6806	99.9	1255	9 US-09-811-123-9	Sequence 3, Appl1
9	6806	99.9	1255	9 US-09-811-115-3	Sequence 9, Appl1
10	6806	99.9	1255	12 US-10-207-498-6	Sequence 6, Appl1
11	6806	99.9	1255	12 US-10-338-730-2	Sequence 2, Appl1
12	6806	99.9	1255	12 US-10-321-892-4	Sequence 4, Appl1
13	6806	99.9	1255	12 US-10-272-437A-28	Sequence 28, Appl1
14	6806	99.9	1255	12 US-10-117-937-594	Sequence 594, Appl1
15	6806	99.9	1255	12 US-10-435-696-36	Sequence 36, Appl1

16	6806	99.9	1255	15 US-10-177-293-126	Sequence 126, Appl1
17	6804	99.8	1253	15 US-10-146-473-72	Sequence 72, Appl1
18	6803	99.8	1255	12 US-10-149-138-4641	Sequence 4641, Appl1
19	6757	99.1	1255	10 US-09-769-508-2	Sequence 2, Appl1
20	6001.5	88.1	1256	10 US-09-854-356-2	Sequence 2, Appl1
21	5997	88.0	1260	10 US-09-870-759-118	Sequence 118, Appl1
22	5997	88.0	1260	10 US-09-751-708A-118	Sequence 118, Appl1
23	5982.5	87.8	1256	10 US-09-854-356-14	Sequence 14, Appl1
24	4900	71.9	919	10 US-09-854-356-6	Sequence 6, Appl1
25	3798	55.7	690	12 US-10-412-804A-11	Sequence 11, Appl1
26	3776	55.4	712	10 US-09-854-356-7	Sequence 7, Appl1
27	3694.5	54.2	720	12 US-10-412-804A-6	Sequence 6, Appl1
28	3691	54.2	715	12 US-10-412-804A-10	Sequence 10, Appl1
29	3641	53.4	685	12 US-10-412-804A-4	Sequence 4, Appl1
30	3628	53.2	653	10 US-09-854-356-3	Sequence 3, Appl1
31	3590	52.7	645	9 US-09-921-161-1	Sequence 1, Appl1
32	3580	52.7	645	12 US-10-608-628-13	Sequence 13, Appl1
33	3590	52.7	645	15 US-10-268-501-13	Sequence 13, Appl1
34	3169	46.5	1210	9 US-09-725-433-2	Sequence 2, Appl1
35	3167	46.5	1210	12 US-10-394-322A-16	Sequence 16, Appl1
36	3110.5	45.6	654	10 US-09-854-356-8	Sequence 8, Appl1
37	3092	45.4	583	10 US-09-930-125-9	Sequence 9, Appl1
38	3092	45.4	587	10 US-09-930-125-8	Sequence 8, Appl1
39	3091	45.4	589	10 US-09-930-125-10	Sequence 10, Appl1
40	3091	45.4	600	10 US-09-930-125-11	Sequence 11, Appl1
41	3054	44.8	791	12 US-10-378-393-11	Sequence 11, Appl1
42	3040	44.6	570	12 US-10-378-393-15	Sequence 15, Appl1
43	3004.5	44.1	1308	12 US-09-940-101-2	Sequence 2, Appl1
44	3004.5	44.1	1308	12 US-10-394-322A-18	Sequence 18, Appl1
45	3004.5	44.1	1308	15 US-10-207-655-47	Sequence 47, Appl1

ALIGNMENTS

```
RESULT 1
US-09-854-356-1
Sequence 1, Application US/09854356
Patent No. US2002017567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Geyser, Dirk
APPLICANT: Corixa Corporation
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT FILING DATE: 2001-05-09
PRIORITY APPLICATION NUMBER: US 09/493,480
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: US 60/117,976
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human HER-2/neu protein
NAME/KEY: DOMAIN
LOCATION: (1) (653)
OTHER INFORMATION: extracellular domain (ECD)
NAME/KEY: DOMAIN
LOCATION: (676) (1255)
OTHER INFORMATION: intracellular domain (ICD)
NAME/KEY: DOMAIN
LOCATION: (990) (1255)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (990) (1048)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD)
```

US-09-854-356-1

Query Match 100.0%; Score 6815; DB 10; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MELAALCRWGLLALALPPGAASVQCTGTDMLRLPASPTHLDMLRHLYOGCQVQGNL 60
DB 1 MELAALCRWGLLALALPPGAASVQCTGTDMLRLPASPTHLDMLRHLYOGCQVQGNL 60
QY 61 ELTYLPTNASTSLFLDIQEVQVYLLAHNOVROVPLORLRIVGTOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASTSLFLDIQEVQVYLLAHNOVROVPLORLRIVGTOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLTRTVACGACARCKGPLPTDCCHEOC 240
DB 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLTRTVACGACARCKGPLPTDCCHEOC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICEHLCPALVTYNTDTFESMPNBEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACHFNHSGICEHLCPALVTYNTDTFESMPNBEGRYTFGASCVTACP 300
QY 301 YNYLSTDVSGCTLVCPPLHNOEVTABDGTQRCCKSKPCAVCYGLGMEHLREVAVTSAN 360
DB 301 YNYLSTDVSGCTLVCPPLHNOEVTABDGTQRCCKSKPCAVCYGLGMEHLREVAVTSAN 360
QY 361 IQSPAGCKKIFGSLAPLPSFDGDPASNTAPLOEOLQVETTELITGYLYISAMPDLP 420
DB 361 IQSPAGCKKIFGSLAPLPSFDGDPASNTAPLOEOLQVETTELITGYLYISAMPDLP 420
QY 421 DLSVFNOLQVIRGRILHNGAYSLTLOGLISWLGRLRELSGSLALIHNTHLCFHTV 480
DB 421 DLSVFNOLQVIRGRILHNGAYSLTLOGLISWLGRLRELSGSLALIHNTHLCFHTV 480
QY 481 PMDOLFRNPHQALLHTANREDEECVGBGLACHQICARHGMCPGPTQCVNCSOPLRGQC 540
DB 481 PMDOLFRNPHQALLHTANREDEECVGBGLACHQICARHGMCPGPTQCVNCSOPLRGQC 540
QY 541 VESCRVLQGLPREVYNAHRLCPHPECOPONGSYTCGFPADQVACAHYKDPFCYARC 600
DB 541 VESCRVLQGLPREVYNAHRLCPHPECOPONGSYTCGFPADQVACAHYKDPFCYARC 600
QY 601 PGGVKPDLSTYMPIMKPEDEGAOCPCINCTHSCVDLDDKGCAPMORASPLTSLISAVG 660
DB 601 PGGVKPDLSTYMPIMKPEDEGAOCPCINCTHSCVDLDDKGCAPMORASPLTSLISAVG 660
QY 661 ILLVVLGVVFGILLIKRQOKIRKTYMRRLLOETELVEPLTPSGAMPNOAOMRLKETEL 720
DB 661 ILLVVLGVVFGILLIKRQOKIRKTYMRRLLOETELVEPLTPSGAMPNOAOMRLKETEL 720
QY 721 RKRKVTGSGAGFYVYKGIWIPDGENVKI PVAIKYLRKENTSPKANKELIDEAYVMAVGSF 780
DB 721 RKRKVTGSGAGFYVYKGIWIPDGENVKI PVAIKYLRKENTSPKANKELIDEAYVMAVGSF 780
QY 781 YVSRLLGICLTSTVQVLTQMLPVGCLLDHRENRGRIGSODLWMCQIAKMSYLDVDR 840
DB 781 YVSRLLGICLTSTVQVLTQMLPVGCLLDHRENRGRIGSODLWMCQIAKMSYLDVDR 840
QY 841 LVHRDLAARVLYKSPNHVKITDGLARLDDIDETEHADGKXPIKMALESILRRRFT 900
DB 841 LVHRDLAARVLYKSPNHVKITDGLARLDDIDETEHADGKXPIKMALESILRRRFT 900
QY 901 HQSDVMSYGVTWELMTFGAKPYDGIIPARBIPLLEKGERLPORPITCIDVYTMIVKCM 960
DB 901 HQSDVMSYGVTWELMTFGAKPYDGIIPARBIPLLEKGERLPORPITCIDVYTMIVKCM 960
QY 961 IDSECRPRFRELVESEFRMADPORPVVITQNEBIDGSPASPLDSTFYRSLIBDDMGDLYDA 1020
DB 961 IDSECRPRFRELVESEFRMADPORPVVITQNEBIDGSPASPLDSTFYRSLIBDDMGDLYDA 1020
```

```
QY 1021 EBYLVPQGFPCPDPAAGAGVHHRSSSTRSGGDLTLGLEPSEERAPRSLABSEG 1080
DB 1021 EBYLVPQGFPCPDPAAGAGVHHRSSSTRSGGDLTLGLEPSEERAPRSLABSEG 1080
QY 1081 ASDVFDGDLGMAAGLOSPLPHDPSPLORVSEDPVLPSESTDGVAFLTCSPOREYV 1140
DB 1081 ASDVFDGDLGMAAGLOSPLPHDPSPLORVSEDPVLPSESTDGVAFLTCSPOREYV 1140
QY 1141 NQPDVAPQPSREBGLPARPAGATLERPKTLPKNGVAVKDVAFGAVENPEYLTPO 1200
DB 1141 NQPDVAPQPSREBGLPARPAGATLERPKTLPKNGVAVKDVAFGAVENPEYLTPO 1200
QY 1201 GGAAPQHPPPAFSPAFDNLVYWDODPPERGAPSTFKGPTAENBYLGLDVPV 1255
DB 1201 GGAAPQHPPPAFSPAFDNLVYWDODPPERGAPSTFKGPTAENBYLGLDVPV 1255
```

RESULT 2

```
US-09-930-125-2
; Sequence 2, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Gheever, Martin A.
; APPLICANT: Roy, Teresa M.
; APPLICANT: Iodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-930-125-2
```

Query Match 100.0%; Score 6815; DB 10; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MELAALCRWGLLALALPPGAASVQCTGTDMLRLPASPTHLDMLRHLYOGCQVQGNL 60
DB 1 MELAALCRWGLLALALPPGAASVQCTGTDMLRLPASPTHLDMLRHLYOGCQVQGNL 60
QY 61 ELTYLPTNASTSLFLDIQEVQVYLLAHNOVROVPLORLRIVGTOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASTSLFLDIQEVQVYLLAHNOVROVPLORLRIVGTOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLTRTVACGACARCKGPLPTDCCHEOC 240
DB 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLTRTVACGACARCKGPLPTDCCHEOC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICEHLCPALVTYNTDTFESMPNBEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACHFNHSGICEHLCPALVTYNTDTFESMPNBEGRYTFGASCVTACP 300
QY 301 YNYLSTDVSGCTLVCPPLHNOEVTABDGTQRCCKSKPCAVCYGLGMEHLREVAVTSAN 360
DB 301 YNYLSTDVSGCTLVCPPLHNOEVTABDGTQRCCKSKPCAVCYGLGMEHLREVAVTSAN 360
QY 361 IQSPAGCKKIFGSLAPLPSFDGDPASNTAPLOEOLQVETTELITGYLYISAMPDLP 420
DB 361 IQSPAGCKKIFGSLAPLPSFDGDPASNTAPLOEOLQVETTELITGYLYISAMPDLP 420
```

```

Db      361 IOEPACCKKI FGSIAFLPESGPDDBASNTAPLOPEOUQVETLEITGYLISAWPDSL 420
Qy      421 DLSVPONTQVIRGRILINNGAYSLTLOGISWLGRLSRLSGSLALIHNTLHCFVHTV 480
Db      421 DLSVPONTQVIRGRILINNGAYSLTLOGISWLGRLSRLSGSLALIHNTLHCFVHTV 480
Qy      481 PMDOLFRNPHOALHTANRPEDECVGEGLAHQULCARGHGWGPGPTCCVNCOSQFLRQEC 540
Db      481 PMDOLFRNPHOALHTANRPEDECVGEGLAHQULCARGHGWGPGPTCCVNCOSQFLRQEC 540
Qy      541 VECRCVLOGLPREYVNAHRLCPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Db      541 VECRCVLOGLPREYVNAHRLCPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Qy      601 PSQVKDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCRAEORASPLTSIIISAVG 660
Db      601 PSQVKDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCRAEORASPLTSIIISAVG 660
Qy      661 ILLVVLGVVFGILLIKRQOKIKRTMRLLQETELVEPLTPSGAMPNOQMILKXETEL 720
Db      661 ILLVVLGVVFGILLIKRQOKIKRTMRLLQETELVEPLTPSGAMPNOQMILKXETEL 720
Qy      721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVLRNTPSPKANKETLDEAYYMAVGSP 780
Db      721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVLRNTPSPKANKETLDEAYYMAVGSP 780
Qy      781 YVSRLLGICLTSTVOLVQMLPYGCLLDHVRNRRGLSODLNMCMQIAKMSYLEDR 840
Db      781 YVSRLLGICLTSTVOLVQMLPYGCLLDHVRNRRGLSODLNMCMQIAKMSYLEDR 840
Qy      841 LVHRDLAANVAVKSPNHVKITDFGLARLLDIDETEHADGKVPKIMWALLESILRRFT 900
Db      841 LVHRDLAANVAVKSPNHVKITDFGLARLLDIDETEHADGKVPKIMWALLESILRRFT 900
Qy      901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGRLLPOPICTIDVYIMVYKCM 960
Db      901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGRLLPOPICTIDVYIMVYKCM 960
Qy      961 IOSECRPRELVESESRMARDPQRFVIONEDLGPASPLDSTFYSLLEDDMDGLVDA 1020
Db      961 IOSECRPRELVESESRMARDPQRFVIONEDLGPASPLDSTFYSLLEDDMDGLVDA 1020
Qy      1021 BEYLVPQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLEPSEBBAFRSPLAPSEG 1080
Db      1021 BEYLVPQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLEPSEBBAFRSPLAPSEG 1080
Qy      1081 ASSDVFDGLGMAAGLQSLPTHDSPLQRYSEDTVPPLPSETDGYVAPLTCSPQPEYV 1140
Db      1081 ASSDVFDGLGMAAGLQSLPTHDSPLQRYSEDTVPPLPSETDGYVAPLTCSPQPEYV 1140
Qy      1141 NOPDVAPQPPPREGFLPARPAGATLERPKLSPKNGVVDVFAFGAVENPEYLTQ 1200
Db      1141 NOPDVAPQPPPREGFLPARPAGATLERPKLSPKNGVVDVFAFGAVENPEYLTQ 1200
Qy      1201 GGAAPQHPHPPAPSPADNLVYWDQDPERKAPSTFGKGTPTAENPEYGLDVPV 1255
Db      1201 GGAAPQHPHPPAPSPADNLVYWDQDPERKAPSTFGKGTPTAENPEYGLDVPV 1255

```

```

; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-441-411-6

Query Match      100.0%; Score 6815; DB 11; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 METALCRMGILLIALLPFGAASQVCTGDMKRLRPAPEYTHLDMRLHYOGCGVOGNTL 60
Db      1 METALCRMGILLIALLPFGAASQVCTGDMKRLRPAPEYTHLDMRLHYOGCGVOGNTL 60
Qy      61 EITVLPFNASLSPLODIQEVQGVVLAHNOVRQVPLQRLATVNGTOLFEDNYALAVDNG 120
Db      61 EITVLPFNASLSPLODIQEVQGVVLAHNOVRQVPLQRLATVNGTOLFEDNYALAVDNG 120
Qy      121 DFLANTTPVTCASPGGLREHQLRSLEILKGVLIQBNPOLCYODTILMKDIFHKNNOLA 180
Db      121 DFLANTTPVTCASPGGLREHQLRSLEILKGVLIQBNPOLCYODTILMKDIFHKNNOLA 180
Qy      181 LTLIDNRSRACHPCSPMCKGSRCWGESSEDCQLTRTVACGAGCARCKGPLPTDCHEOC 240
Db      181 LTLIDNRSRACHPCSPMCKGSRCWGESSEDCQLTRTVACGAGCARCKGPLPTDCHEOC 240
Qy      241 AAGCTGPKSDCLACIHPNHSIGCEIHCPLVNTYNTDFESMNPBGRYTFGASCVTACP 300
Db      241 AAGCTGPKSDCLACIHPNHSIGCEIHCPLVNTYNTDFESMNPBGRYTFGASCVTACP 300
Qy      301 YVLTSTVGSCTVCPILHNOVTAEDGTORCEKSCRPACAVCYGLGMEHLREVAATVSAN 360
Db      301 YVLTSTVGSCTVCPILHNOVTAEDGTORCEKSCRPACAVCYGLGMEHLREVAATVSAN 360
Qy      361 IOEPACCKKI FGSIAFLPESGPDDBASNTAPLOPEOUQVETLEITGYLISAWPDSL 420
Db      361 IOEPACCKKI FGSIAFLPESGPDDBASNTAPLOPEOUQVETLEITGYLISAWPDSL 420
Qy      421 DLSVPONTQVIRGRILINNGAYSLTLOGISWLGRLSRLSGSLALIHNTLHCFVHTV 480
Db      421 DLSVPONTQVIRGRILINNGAYSLTLOGISWLGRLSRLSGSLALIHNTLHCFVHTV 480
Qy      481 PMDOLFRNPHOALHTANRPEDECVGEGLAHQULCARGHGWGPGPTCCVNCOSQFLRQEC 540
Db      481 PMDOLFRNPHOALHTANRPEDECVGEGLAHQULCARGHGWGPGPTCCVNCOSQFLRQEC 540
Qy      541 VECRCVLOGLPREYVNAHRLCPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Db      541 VECRCVLOGLPREYVNAHRLCPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Qy      601 PSQVKDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCRAEORASPLTSIIISAVG 660
Db      601 PSQVKDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDKGCRAEORASPLTSIIISAVG 660
Qy      661 ILLVVLGVVFGILLIKRQOKIKRTMRLLQETELVEPLTPSGAMPNOQMILKXETEL 720
Db      661 ILLVVLGVVFGILLIKRQOKIKRTMRLLQETELVEPLTPSGAMPNOQMILKXETEL 720
Qy      721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVLRNTPSPKANKETLDEAYYMAVGSP 780
Db      721 RKVKVLGSGAFGVYKGIWIPDGENVKIPVAIKVLRNTPSPKANKETLDEAYYMAVGSP 780
Qy      781 YVSRLLGICLTSTVOLVQMLPYGCLLDHVRNRRGLSODLNMCMQIAKMSYLEDR 840
Db      781 YVSRLLGICLTSTVOLVQMLPYGCLLDHVRNRRGLSODLNMCMQIAKMSYLEDR 840
Qy      841 LVHRDLAANVAVKSPNHVKITDFGLARLLDIDETEHADGKVPKIMWALLESILRRFT 900
Db      841 LVHRDLAANVAVKSPNHVKITDFGLARLLDIDETEHADGKVPKIMWALLESILRRFT 900
Qy      901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGRLLPOPICTIDVYIMVYKCM 960
Db      901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGRLLPOPICTIDVYIMVYKCM 960

```

```

RESULT 3
US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Diels, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26

```

Db 901 HOSDWSYGVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPPITIDVYIMVKCM 960
Qy 961 IDSECRPRFRELSEFSRMARDPQRFVYIIONEDLGPAAPLSTFYRSILBDDMDGLVDA 1020
Db 961 IDSECRPRFRELSEFSRMARDPQRFVYIIONEDLGPAAPLSTFYRSILBDDMDGLVDA 1020
Qy 1021 EBYLVPOQGFCCDPAPAGAGMYHHRSSSTSSGGGDLTLGLPSESEAPRSLAPSEG 1080
Db 1021 EBYLVPOQGFCCDPAPAGAGMYHHRSSSTSSGGGDLTLGLPSESEAPRSLAPSEG 1080
Qy 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLORYSEDPVPLPSENDGVVAPLTCSPOPEYV 1140
Db 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLORYSEDPVPLPSENDGVVAPLTCSPOPEYV 1140
Qy 1141 NOPDVPRPSPREGPLPARPAGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
Db 1141 NOPDVPRPSPREGPLPARPAGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPERGAPSTKGTPTAENPEYLGIDVAV 1255
Db 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPERGAPSTKGTPTAENPEYLGIDVAV 1255

RESULT 4
US-10-313-644-2
Sequence 2, Application US/10313644
Publication No. US20030157119A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Cheever, Martin A.
APPLICANT: Hand-Zimmerman, Susan
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
TITLE OF INVENTION: AND VIRUS-ASSOCIATED MALIGNANCIES
FILE REFERENCE: 210121.483C3
CURRENT APPLICATION NUMBER: US/10/313.644
CURRENT FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 3.0
LENGTH: 1255
SEQ ID NO: 2
TYPE: PRN
ORGANISM: Homo. sapien
US-10-313-644-2

Query Match 100.0%; Score 6815; DB 12; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEIAALCRWGLLALLPFGAASSTVCTGTDMKRLPASPETHLDMLEHLYQGCQVQGNL 60
Db 1 MEIAALCRWGLLALLPFGAASSTVCTGTDMKRLPASPETHLDMLEHLYQGCQVQGNL 60
Qy 61 ELYTLPNABLSPLQDIQEVQGYVLIANOVROVPLQRLRVKRTQVPEYNYALAVDNG 120
Db 61 ELYTLPNABLSPLQDIQEVQGYVLIANOVROVPLQRLRVKRTQVPEYNYALAVDNG 120
Qy 121 DPLANTTPVYVAGSGGLRELOLRLTLKLGVLIOBNPOLCYODTILMKDIFKKNOLA 180
Db 121 DPLANTTPVYVAGSGGLRELOLRLTLKLGVLIOBNPOLCYODTILMKDIFKKNOLA 180
Qy 181 LTLIDTRSRACHPCSPWCKGSRGSSSBDQSLTRTVCAAGCARCKGPLPTDCHEQC 240
Db 181 LTLIDTRSRACHPCSPWCKGSRGSSSBDQSLTRTVCAAGCARCKGPLPTDCHEQC 240
Qy 241 AAGCTGKSHDCLALPHNHSIGIELCPALVYNTOTFESMNPBERGRTYVAGASCYVACP 300
Db 241 AAGCTGKSHDCLALPHNHSIGIELCPALVYNTOTFESMNPBERGRTYVAGASCYVACP 300
Qy 301 YNLTSTVGSCTVLCPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVAAYTSAN 360
Db 301 YNLTSTVGSCTVLCPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVAAYTSAN 360
Qy 361 IOBPAGCKKIFGLAFLPESFSDGPASNTAPLOPEOLQVFTLEETIGYLYISAMPDLP 420

Db 361 IOBPAGCKKIFGLAFLPESFSDGPASNTAPLOPEOLQVFTLEETIGYLYISAMPDLP 420
Qy 421 DLSVFQNLQVIRIRILHNGAYSLTLQGLISWIGLSRLRELSGALILHNTHLCPVHTV 480
Db 421 DLSVFQNLQVIRIRILHNGAYSLTLQGLISWIGLSRLRELSGALILHNTHLCPVHTV 480
Qy 481 PMDOLFNNPHQALLHTANRPEDECEVSEGLACHOLCARGHCWGGPTQCVNCSQPLRQDEC 540
Db 481 PMDOLFNNPHQALLHTANRPEDECEVSEGLACHOLCARGHCWGGPTQCVNCSQPLRQDEC 540
Qy 541 VEECRVLQGLPREYVNAARHCLPCHPCQPNQSVTCFEPADQCVAAHKKDPFCVARC 600
Db 541 VEECRVLQGLPREYVNAARHCLPCHPCQPNQSVTCFEPADQCVAAHKKDPFCVARC 600
Qy 601 PEGVPEDLSTMPYMKPDEBGACQPCPINCTHSCVDLDDKGCFAEQASPLTSIIISAVG 660
Db 601 PEGVPEDLSTMPYMKPDEBGACQPCPINCTHSCVDLDDKGCFAEQASPLTSIIISAVG 660
Qy 661 ILTVVLGVVFGILIRROOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRLKETEL 720
Db 661 ILTVVLGVVFGILIRROOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRLKETEL 720
Qy 721 RKRYUGSAGFQVYVGIWTPDEGNVKIPIAIVLRENTSPKANKEILDEAYVVAAGVSP 780
Db 721 RKRYUGSAGFQVYVGIWTPDEGNVKIPIAIVLRENTSPKANKEILDEAYVVAAGVSP 780
Qy 781 YVSRLLGICLSTVQVLTQMLPYGCLLDHYRENGRLGSDLLNMCQIAKMSYLEBVR 840
Db 781 YVSRLLGICLSTVQVLTQMLPYGCLLDHYRENGRLGSDLLNMCQIAKMSYLEBVR 840
Qy 841 LVHRDLAARVAVKSPNHVKITDFGLARLIDIDETEHADGKVPYKMALESILRRFT 900
Db 841 LVHRDLAARVAVKSPNHVKITDFGLARLIDIDETEHADGKVPYKMALESILRRFT 900
Qy 901 HOSDWSYGVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPPITIDVYIMVKCM 960
Db 901 HOSDWSYGVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPPITIDVYIMVKCM 960
Qy 961 IDSECRPRFRELSEFSRMARDPQRFVYIIONEDLGPAAPLSTFYRSILBDDMDGLVDA 1020
Db 961 IDSECRPRFRELSEFSRMARDPQRFVYIIONEDLGPAAPLSTFYRSILBDDMDGLVDA 1020
Qy 1021 EBYLVPOQGFCCDPAPAGAGMYHHRSSSTSSGGGDLTLGLPSESEAPRSLAPSEG 1080
Db 1021 EBYLVPOQGFCCDPAPAGAGMYHHRSSSTSSGGGDLTLGLPSESEAPRSLAPSEG 1080
Qy 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLORYSEDPVPLPSENDGVVAPLTCSPOPEYV 1140
Db 1081 AGSDVDGDLGMAAGLQSLPTHDPSPLORYSEDPVPLPSENDGVVAPLTCSPOPEYV 1140
Qy 1141 NOPDVPRPSPREGPLPARPAGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
Db 1141 NOPDVPRPSPREGPLPARPAGATLERPKTSLPGKNGVYKDFAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPERGAPSTKGTPTAENPEYLGIDVAV 1255
Db 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPERGAPSTKGTPTAENPEYLGIDVAV 1255

RESULT 5
US-10-394-322A-17
Sequence 17, Application US/10394322A
Publication No. US20030232391A1
GENERAL INFORMATION:
APPLICANT: SYNESIS PHARMACEUTICALS, INC.
APPLICANT: Prescott, John C.
TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
FILE REFERENCE: 39750-0006 US
CURRENT APPLICATION NUMBER: US/10/394.322A
CURRENT FILING DATE: 2003-03-20
PRIOR APPLICATION NUMBER: US 60/366,892
PRIOR FILING DATE: 2002-03-21

NUMBER OF SEQ ID NOS: 70
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 17
 LENGTH: 1255
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-394-322A-17

Query Match 100.0%; Score 6815; DB 12; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELAALCRWGLIALLPFGAASSTVCTGDMKRLPASPEHLMRLHLYOGCQVVGNTL 60
DB 1 MELAALCRWGLIALLPFGAASSTVCTGDMKRLPASPEHLMRLHLYOGCQVVGNTL 60
QY 61 ELTYLPTNASLSPLODIOEVGVYLIANOVQVPLRLIRVGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSPLODIOEVGVYLIANOVQVPLRLIRVGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
DB 121 DPLNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
QY 181 LTLIDTNSRACHPCSPCKGSRCWGSSSDCOSLTRTVCAAGCARKGPLPTDCHEOC 240
DB 181 LTLIDTNSRACHPCSPCKGSRCWGSSSDCOSLTRTVCAAGCARKGPLPTDCHEOC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPREGRTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPREGRTFGASCVTACP 300
QY 301 YNTLSTVGSCTLYVCPHNOETVADGTORCEKSKPCAVCYGLGMEHLREVAVTSAN 360
DB 301 YNTLSTVGSCTLYVCPHNOETVADGTORCEKSKPCAVCYGLGMEHLREVAVTSAN 360
QY 361 IOEPACKKIFGSLAPLPSFDGDPASNTAPLOPELQVETLEETGYLYISAMPDSL 420
DB 361 IOEPACKKIFGSLAPLPSFDGDPASNTAPLOPELQVETLEETGYLYISAMPDSL 420
QY 421 DLSVPNTLQVIRGRIHNGAYSLTLQGLGTSWLGSLRLSGSLALIHNTHLCPHTV 480
DB 421 DLSVPNTLQVIRGRIHNGAYSLTLQGLGTSWLGSLRLSGSLALIHNTHLCPHTV 480
QY 481 PMDQLFRNPHOALHTANPEDECVSGEGACHOLCARHGMGPGPTQCVNCSQPLRGQEC 540
DB 481 PMDQLFRNPHOALHTANPEDECVSGEGACHOLCARHGMGPGPTQCVNCSQPLRGQEC 540
QY 541 VEECRVLOGLPRRYVNAHCLPCHPCQPNQSVTCFGEPADQCVACAHYKDPFCVARC 600
DB 541 VEECRVLOGLPRRYVNAHCLPCHPCQPNQSVTCFGEPADQCVACAHYKDPFCVARC 600
QY 601 PEGVPRDLSYMPIMKPEDEGACOPRINCTHS CVLDDKGCABRASPSTIISAVNG 660
DB 601 PEGVPRDLSYMPIMKPEDEGACOPRINCTHS CVLDDKGCABRASPSTIISAVNG 660
QY 661 ILVVVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPNOAKRIKETEL 720
DB 661 ILVVVVLGVVFGILIKRROOKIRKTYMRLLQETELVEPLTPSGAMPNOAKRIKETEL 720
QY 721 RRVKVLGSGAFGTIVKGIWIPGSENYKIVALKVLENTSPKANKIILDEAYMAGVSP 780
DB 721 RRVKVLGSGAFGTIVKGIWIPGSENYKIVALKVLENTSPKANKIILDEAYMAGVSP 780
QY 781 YVSRLLGICLTSTVOLVTLMPYGLLDHVRNRGLSGODLLNMCMQAKMSITIEDVR 840
DB 781 YVSRLLGICLTSTVOLVTLMPYGLLDHVRNRGLSGODLLNMCMQAKMSITIEDVR 840
QY 841 LVHRDLAARNVAVKSPNHXYITDIFGLARLLDIDETFEHADGGKVPFKMALSSILRRRT 900
DB 841 LVHRDLAARNVAVKSPNHXYITDIFGLARLLDIDETFEHADGGKVPFKMALSSILRRRT 900
QY 901 HOSDWSIGVTVWELMTFGAKPYDGIIPAREIPDLLKEGRLLPQPICTIDVYMINVKCM 960
DB 901 HOSDWSIGVTVWELMTFGAKPYDGIIPAREIPDLLKEGRLLPQPICTIDVYMINVKCM 960

```

```

DB 901 HOSDWSIGVTVWELMTFGAKPYDGIIPAREIPDLLKEGRLLPQPICTIDVYMINVKCM 960
QY 961 IDSECRPRPRELVESEFARMADPQRFVITQNEIDGPRASPLSTFRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRPRELVESEFARMADPQRFVITQNEIDGPRASPLSTFRSLLEDDMDGLVDA 1020
QY 1021 EBYLVPOOGFPCPDPAAGAGVHHRRSSSTRSGGGDLTGLBPESEAPRSPAPSEG 1080
DB 1021 EBYLVPOOGFPCPDPAAGAGVHHRRSSSTRSGGGDLTGLBPESEAPRSPAPSEG 1080
QY 1081 AGSDVFPDGLGMAKGLQSLPTHDPSPLQKSSSDPTVPLPSETDGVYAPLTCSPQEV 1140
DB 1081 AGSDVFPDGLGMAKGLQSLPTHDPSPLQKSSSDPTVPLPSETDGVYAPLTCSPQEV 1140
QY 1141 NOPDVRPOPSPREGPLPAPRAGATLERPKTSLPGKNGVYKDVAFAGAVENPEYLTPQ 1200
DB 1141 NOPDVRPOPSPREGPLPAPRAGATLERPKTSLPGKNGVYKDVAFAGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPPPAPSPAFDNLTYMDQDPERGAPSTFGPTAENPEYLGLDVPV 1255
DB 1201 GGAAPQHPPPAPSPAFDNLTYMDQDPERGAPSTFGPTAENPEYLGLDVPV 1255

```

RESULT 6

```

US-10-245-871-553
; Sequence 553, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPI TOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 553
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-245-871-553

```

Query Match 100.0%; Score 6815; DB 12; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELAALCRWGLIALLPFGAASSTVCTGDMKRLPASPEHLMRLHLYOGCQVVGNTL 60
DB 1 MELAALCRWGLIALLPFGAASSTVCTGDMKRLPASPEHLMRLHLYOGCQVVGNTL 60
QY 61 ELTYLPTNASLSPLODIOEVGVYLIANOVQVPLRLIRVGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSPLODIOEVGVYLIANOVQVPLRLIRVGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
DB 121 DPLNNTTPVTGASPGGLRELQLSLTELKGVLIQSNPOLCYQDTILMKDIFHKNQLA 180
QY 181 LTLIDTNSRACHPCSPCKGSRCWGSSSDCOSLTRTVCAAGCARKGPLPTDCHEOC 240
DB 181 LTLIDTNSRACHPCSPCKGSRCWGSSSDCOSLTRTVCAAGCARKGPLPTDCHEOC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPREGRTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPREGRTFGASCVTACP 300
QY 301 YNTLSTVGSCTLYVCPHNOETVADGTORCEKSKPCAVCYGLGMEHLREVAVTSAN 360
DB 301 YNTLSTVGSCTLYVCPHNOETVADGTORCEKSKPCAVCYGLGMEHLREVAVTSAN 360

```

```

Db      301 YNTLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKCSKPCARVCYGLGMEHLREVAVTSAN 360
Qy      361 IOEFACCKKI FGSIAFLPESFDDDPASNTAPLOPEOLQVETLEITGYLYISAMPDSLP 420
        361 IOEFACCKKI FGSIAFLPESFDDDPASNTAPLOPEOLQVETLEITGYLYISAMPDSLP 420
Qy      421 DLSVPONLOVIRGRILHNGAYSLTLQGLGISWLGSLNLSRELGSGLAIHHNTHLCFVHTV 480
        421 DLSVPONLOVIRGRILHNGAYSLTLQGLGISWLGSLNLSRELGSGLAIHHNTHLCFVHTV 480
Db      481 PMOQLFRNPHQALLHTANRPEDCEVGEGLACHOLCARGHCWGPGPTOCVNSQFLRGQEC 540
        481 PMOQLFRNPHQALLHTANRPEDCEVGEGLACHOLCARGHCWGPGPTOCVNSQFLRGQEC 540
Qy      541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
        541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Db      601 PSQVKPDLSTMPYIMKPFDEGACOPCPINCTHSCVDLDDKGCAPAEORASPLTISIISAVG 660
        601 PSQVKPDLSTMPYIMKPFDEGACOPCPINCTHSCVDLDDKGCAPAEORASPLTISIISAVG 660
Qy      661 ILLVVLGVVFGILIKRQOKIRKRYTMRLLQETELVEPLTPSGAMNOAMRLKETEL 720
        661 ILLVVLGVVFGILIKRQOKIRKRYTMRLLQETELVEPLTPSGAMNOAMRLKETEL 720
Db      721 RKRYVLGSAGFGVYKGIWI PDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
        721 RKRYVLGSAGFGVYKGIWI PDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
Qy      781 YSRLLGICLSTVQVLTOLMPYGCCLDHYRENGRGLSGODLNMCMQIAKMSYLEDR 840
        781 YSRLLGICLSTVQVLTOLMPYGCCLDHYRENGRGLSGODLNMCMQIAKMSYLEDR 840
Db      841 LVHNDLAARAVLVSPNHVKITDGLARLLDIDETEHADGKVPKMMALESIIRRRFT 900
        841 LVHNDLAARAVLVSPNHVKITDGLARLLDIDETEHADGKVPKMMALESIIRRRFT 900
Qy      901 HOSDWSYGVYVWELMTFGAKFYDGI PABEIPDLLEKGERLPOPPICITIDVYIMVCKM 960
        901 HOSDWSYGVYVWELMTFGAKFYDGI PABEIPDLLEKGERLPOPPICITIDVYIMVCKM 960
Db      961 IDSECRPRFELVSEFSRMARDPFRVYVITONEDIGRASPDLSTYRSLLEBDDMDGLVDA 1020
        961 IDSECRPRFELVSEFSRMARDPFRVYVITONEDIGRASPDLSTYRSLLEBDDMDGLVDA 1020
Qy      1021 BEYVLPQOQFPCPDPAAGAGMHHRRHSSSTRSGGDLTLGLPSEBEARSPFLASSEG 1080
        1021 BEYVLPQOQFPCPDPAAGAGMHHRRHSSSTRSGGDLTLGLPSEBEARSPFLASSEG 1080
Db      1081 AGSDVPDGLCMGAAGKIQSLPTHDSPLOHSEDPVPLPSETDGVAVPLTCSPOBEYV 1140
        1081 AGSDVPDGLCMGAAGKIQSLPTHDSPLOHSEDPVPLPSETDGVAVPLTCSPOBEYV 1140
Qy      1141 NQPDVPRQPPSPRSGPLPAPAPAGATLERPTLSFGKGVVYKDVAFAGGAVENBEYLPQ 1200
        1141 NQPDVPRQPPSPRSGPLPAPAPAGATLERPTLSFGKGVVYKDVAFAGGAVENBEYLPQ 1200
Db      1201 GGAAPOHPHPAPAFAPNLYYMDODPPEKAPSTFKGTPTANPEYLGIDVY 1255
        1201 GGAAPOHPHPAPAFAPNLYYMDODPPEKAPSTFKGTPTANPEYLGIDVY 1255

```

RESULT 7

```

US-10-207-655-45
; Sequence 45, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655

```

```

; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 45
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-45

Query Match      100.0%; Score 6815; DB 15; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MEALACRWGLLALLPRAASVQVCTGDMRLRLPASERTHDMRLHYOGQOVQGNL 60
        1 MEALACRWGLLALLPRAASVQVCTGDMRLRLPASERTHDMRLHYOGQOVQGNL 60
Db      61 ELTYLPTNASLFLQDIQEVQGVYLIHNGVQVPLQRLRIYAGTQLFEDNYALAVLNG 120
        61 ELTYLPTNASLFLQDIQEVQGVYLIHNGVQVPLQRLRIYAGTQLFEDNYALAVLNG 120
Qy      121 DELNNTTPTVQASPGGLRELOSLTEILKGVLIQRPOLCYQDTTLMKDIFHKNOQA 180
        121 DELNNTTPTVQASPGGLRELOSLTEILKGVLIQRPOLCYQDTTLMKDIFHKNOQA 180
Db      181 LTIIDTRNSRACHPCSPCKGSRCKWGESSEBDOSLTRTVAGGCARCKGPLPDCCHEOC 240
        181 LTIIDTRNSRACHPCSPCKGSRCKWGESSEBDOSLTRTVAGGCARCKGPLPDCCHEOC 240
Qy      241 AGCTGPKSDCLACAFHNSGICELHCPALVYNTDTEFSMNPBEGRYTFGASCVTACP 300
        241 AGCTGPKSDCLACAFHNSGICELHCPALVYNTDTEFSMNPBEGRYTFGASCVTACP 300
Db      301 YNTLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKCSKPCARVCYGLGMEHLREVAVTSAN 360
        301 YNTLSTDVGSCTLVCPPLHNOEVTAEADGTORCEKCSKPCARVCYGLGMEHLREVAVTSAN 360
Qy      361 IOEFACCKKI FGSIAFLPESFDDDPASNTAPLOPEOLQVETLEITGYLYISAMPDSLP 420
        361 IOEFACCKKI FGSIAFLPESFDDDPASNTAPLOPEOLQVETLEITGYLYISAMPDSLP 420
Db      421 DLSVPONLOVIRGRILHNGAYSLTLQGLGISWLGSLNLSRELGSGLAIHHNTHLCFVHTV 480
        421 DLSVPONLOVIRGRILHNGAYSLTLQGLGISWLGSLNLSRELGSGLAIHHNTHLCFVHTV 480
Qy      481 PMOQLFRNPHQALLHTANRPEDCEVGEGLACHOLCARGHCWGPGPTOCVNSQFLRGQEC 540
        481 PMOQLFRNPHQALLHTANRPEDCEVGEGLACHOLCARGHCWGPGPTOCVNSQFLRGQEC 540
Db      541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
        541 VEECRVLOGLPREYVNAHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVARC 600
Qy      601 PSQVKPDLSTMPYIMKPFDEGACOPCPINCTHSCVDLDDKGCAPAEORASPLTISIISAVG 660
        601 PSQVKPDLSTMPYIMKPFDEGACOPCPINCTHSCVDLDDKGCAPAEORASPLTISIISAVG 660
Db      661 ILLVVLGVVFGILIKRQOKIRKRYTMRLLQETELVEPLTPSGAMNOAMRLKETEL 720
        661 ILLVVLGVVFGILIKRQOKIRKRYTMRLLQETELVEPLTPSGAMNOAMRLKETEL 720
Qy      721 RKRYVLGSAGFGVYKGIWI PDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
        721 RKRYVLGSAGFGVYKGIWI PDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
Db      781 YSRLLGICLSTVQVLTOLMPYGCCLDHYRENGRGLSGODLNMCMQIAKMSYLEDR 840
        781 YSRLLGICLSTVQVLTOLMPYGCCLDHYRENGRGLSGODLNMCMQIAKMSYLEDR 840
Qy      841 LVHNDLAARAVLVSPNHVKITDGLARLLDIDETEHADGKVPKMMALESIIRRRFT 900
        841 LVHNDLAARAVLVSPNHVKITDGLARLLDIDETEHADGKVPKMMALESIIRRRFT 900

```

Qy 901 HOSDWSYGVTVWELMTFGAKPYDGI PARBI PDLLEKGBRLPOPRTCTIDVYMIWKCWM 960
 Db 901 HOSDWSYGVTVWELMTFGAKPYDGI PARBI PDLLEKGBRLPOPRTCTIDVYMIWKCWM 960
 Qy 961 IDSECRPRFRELVSFSSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLBEDDMGDLVDA 1020
 Db 961 IDSECRPRFRELVSFSSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLBEDDMGDLVDA 1020
 Qy 1021 BEVLVPOQGFPCPDPAFGAGGVVHRRSSSTRSGGDLTLGLEPSEEBAPRSLPASEG 1080
 Db 1021 BEVLVPOQGFPCPDPAFGAGGVVHRRSSSTRSGGDLTLGLEPSEEBAPRSLPASEG 1080
 Qy 1081 AGSDVDPDGLGMAAGLQSLPTHDPSPLOYSBDPTVPLPSETDGVVALTCSPOPEYV 1140
 Db 1081 AGSDVDPDGLGMAAGLQSLPTHDPSPLOYSBDPTVPLPSETDGVVALTCSPOPEYV 1140
 Qy 1141 NOPDVAPQPPSPREGLPARPAGATLERPKTSLPGKNGVAVDFAGAVENPEYLTPO 1200
 Db 1141 NOPDVAPQPPSPREGLPARPAGATLERPKTSLPGKNGVAVDFAGAVENPEYLTPO 1200
 Qy 1201 GGAAPQPPPPAPSPAFDNLVYWDOPPERGAPSTFKGTPTAENPEYLGLDVAV 1255
 Db 1201 GGAAPQPPPPAPSPAFDNLVYWDOPPERGAPSTFKGTPTAENPEYLGLDVAV 1255

RESULT 8

US-09-811-123-9
 / Sequence 9, Application US/09811123
 / Patent No. US20020001587A1
 / GENERAL INFORMATION:
 / APPLICANT: Sharon Erickson
 / APPLICANT: Ralph Schwall
 / APPLICANT: Mark Sliwowski
 / TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EBB
 / FILE REFERENCE: GENENT. 073A2
 / CURRENT APPLICATION NUMBER: US/09/811,123
 / PRIOR FILING DATE: 2001-03-16
 / PRIOR APPLICATION NUMBER: 60/238,327
 / PRIOR FILING DATE: 2000-10-05
 / PRIOR APPLICATION NUMBER: 09/602,530
 / NUMBER OF SEQ ID NOS: 11
 / SOFTWARE: Faetsco for Windows Version 4.0
 / SEQ ID NO 9
 / LENGTH: 1255
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-811-123-9

Query Match 99.9%; Score 6806; DB 9; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELALCCKWGLLALIPPAASTOYCTGTDMKRLPASPETHLDMRLHYGCGQVVGNTL 60
 Db 1 MELALCCKWGLLALIPPAASTOYCTGTDMKRLPASPETHLDMRLHYGCGQVVGNTL 60
 Qy 61 ELTYLPTNASLSFLDIOVOGVYLAHQVQVPLRLRIYRGQLFEDNALAVLDNG 120
 Db 61 ELTYLPTNASLSFLDIOVOGVYLAHQVQVPLRLRIYRGQLFEDNALAVLDNG 120
 Qy 121 DPLNNTTPVTGASPGGLRELQLRSLEILKGVLLIQRNQLCYODITILMKDIFHKNNOLA 180
 Db 121 DPLNNTTPVTGASPGGLRELQLRSLEILKGVLLIQRNQLCYODITILMKDIFHKNNOLA 180
 Qy 181 LTLIDTNSRACHPSCSPKCKSGRCSGSESDCQSLTRTYCAGGACARCKGRLPTDCCHQC 240
 Db 181 LTLIDTNSRACHPSCSPKCKSGRCSGSESDCQSLTRTYCAGGACARCKGRLPTDCCHQC 240
 Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTNTDTPESMPNBSGRTTBSACTYACP 300
 Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTNTDTPESMPNBSGRTTBSACTYACP 300

Qy 301 YNLSIDVGSCTLVCPHLNQEVTAEADGTORCEKSKRCARVCYGLGMEHLREVAATVSAN 360
 Db 301 YNLSIDVGSCTLVCPHLNQEVTAEADGTORCEKSKRCARVCYGLGMEHLREVAATVSAN 360
 Qy 361 IDFPACCKIFGSLATLPSFDGDPASNTAPLOPEOLQVETLEITGYLYISAMPDLP 420
 Db 361 IDFPACCKIFGSLATLPSFDGDPASNTAPLOPEOLQVETLEITGYLYISAMPDLP 420
 Qy 421 DISFONLOVIRIRILHNGAYSLTLQGLISWGLSLRELSGGLALIHNTHLCPHTV 480
 Db 421 DISFONLOVIRIRILHNGAYSLTLQGLISWGLSLRELSGGLALIHNTHLCPHTV 480
 Qy 481 PMDQLFRNPHOALLHTANRPEDECVGEGLAHQLCARGHCWGPPTQCVNCSQFLRQEC 540
 Db 481 PMDQLFRNPHOALLHTANRPEDECVGEGLAHQLCARGHCWGPPTQCVNCSQFLRQEC 540
 Qy 541 VBERCVLQGLPRRYNARHCLCPHECOPONGSVTCRGPADCCVACAHYKDPFCVARC 600
 Db 541 VBERCVLQGLPRRYNARHCLCPHECOPONGSVTCRGPADCCVACAHYKDPFCVARC 600
 Qy 601 PEGVXKPLSYMPIMKRPDEEGACOPCINCTHSQVLDLDDGCPAEQASPLTISAVVG 660
 Db 601 PEGVXKPLSYMPIMKRPDEEGACOPCINCTHSQVLDLDDGCPAEQASPLTISAVVG 660
 Qy 661 ILLVVLGVVFGILLKROOKIRKTYMRLLQETELVEPLTPSGAMPNQAOMRLKETEL 720
 Db 661 ILLVVLGVVFGILLKROOKIRKTYMRLLQETELVEPLTPSGAMPNQAOMRLKETEL 720
 Qy 721 RKXKVLGSGAFGTIVKGIWIPGENYKIPVAIKVIRENTSPXANKELDEAYYAVAGVSP 780
 Db 721 RKXKVLGSGAFGTIVKGIWIPGENYKIPVAIKVIRENTSPXANKELDEAYYAVAGVSP 780
 Qy 781 YSRRLIGLSTSTVOLVQMLMPYGLLDHVRNRGSLGODILNMCQIAKGSYLEDVR 840
 Db 781 YSRRLIGLSTSTVOLVQMLMPYGLLDHVRNRGSLGODILNMCQIAKGSYLEDVR 840
 Qy 841 LVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETRYHADGGKVPKMMALSSILRRFT 900
 Db 841 LVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETRYHADGGKVPKMMALSSILRRFT 900
 Qy 901 HOSDWSYGVTVWELMTFGAKPYDGI PARBI PDLLEKGBRLPOPRTCTIDVYMIWKCWM 960
 Db 901 HOSDWSYGVTVWELMTFGAKPYDGI PARBI PDLLEKGBRLPOPRTCTIDVYMIWKCWM 960
 Qy 961 IDSECRPRFRELVSFSSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLBEDDMGDLVDA 1020
 Db 961 IDSECRPRFRELVSFSSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLBEDDMGDLVDA 1020
 Qy 1021 BEVLVPOQGFPCPDPAFGAGGVVHRRSSSTRSGGDLTLGLEPSEEBAPRSLPASEG 1080
 Db 1021 BEVLVPOQGFPCPDPAFGAGGVVHRRSSSTRSGGDLTLGLEPSEEBAPRSLPASEG 1080
 Qy 1081 AGSDVDPDGLGMAAGLQSLPTHDPSPLOYSBDPTVPLPSETDGVVALTCSPOPEYV 1140
 Db 1081 AGSDVDPDGLGMAAGLQSLPTHDPSPLOYSBDPTVPLPSETDGVVALTCSPOPEYV 1140
 Qy 1141 NOPDVAPQPPSPREGLPARPAGATLERPKTSLPGKNGVAVDFAGAVENPEYLTPO 1200
 Db 1141 NOPDVAPQPPSPREGLPARPAGATLERPKTSLPGKNGVAVDFAGAVENPEYLTPO 1200
 Qy 1201 GGAAPQPPPPAPSPAFDNLVYWDOPPERGAPSTFKGTPTAENPEYLGLDVAV 1255
 Db 1201 GGAAPQPPPPAPSPAFDNLVYWDOPPERGAPSTFKGTPTAENPEYLGLDVAV 1255

RESULT 9

US-09-811-115-3
 / Sequence 3, Application US/09811115
 / Patent No. US2002003576A1
 / GENERAL INFORMATION:
 / APPLICANT: Erickson, Sharon
 / APPLICANT: Schwall, Ralph

```
APPLICANT: King, Kathleen
TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
FILE REFERENCE: GENENT. 034A
CURRENT APPLICATION NUMBER: US/09/811, 115
CURRENT FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/189, 844
PRIOR FILING DATE: 2000-03-16
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-09-811-115-3

Query Match      99.9%; Score 6806; DB 9; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  MELAALCRWGLLALLPFGAASSTOVCTGTMKRLRPSPETHLDMRLHLYOGCOVQGNL 60
Db      1  MELAALCRWGLLALLPFGAASSTOVCTGTMKRLRPSPETHLDMRLHLYOGCOVQGNL 60

Qy      61  ELTYLPTNASLSFLQDIQEVGYVLIHNOVROVPLRLRIVRGTOLEFEDNYALAVLDNG 120
Db      61  ELTYLPTNASLSFLQDIQEVGYVLIHNOVROVPLRLRIVRGTOLEFEDNYALAVLDNG 120

Qy      121  DPLANTPTVPGASPGGRLQLSLRTEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180
Db      121  DPLANTPTVPGASPGGRLQLSLRTEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180

Qy      121  DPLANTPTVPGASPGGRLQLSLRTEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180
Db      121  DPLANTPTVPGASPGGRLQLSLRTEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180

Qy      181  LTLIDNRSRAKCPSCMCKGSRCKWSSSDCOSLTTTVACGACGACRCKPLPTDCCHQC 240
Db      181  LTLIDNRSRAKCPSCMCKGSRCKWSSSDCOSLTTTVACGACGACRCKPLPTDCCHQC 240

Qy      181  LTLIDNRSRAKCPSCMCKGSRCKWSSSDCOSLTTTVACGACGACRCKPLPTDCCHQC 240
Db      181  LTLIDNRSRAKCPSCMCKGSRCKWSSSDCOSLTTTVACGACGACRCKPLPTDCCHQC 240

Qy      241  AACCTGPKSHDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCTYACP 300
Db      241  AACCTGPKSHDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCTYACP 300

Qy      241  AACCTGPKSHDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCTYACP 300
Db      241  AACCTGPKSHDCLACLFHNSGICELHCPALVTYNTDFESMNPBGRYTFGASCTYACP 300

Qy      301  YNLTSDVGSCTLVCPILHNOEVTAEQTCCKSKCAVCVGLGMEHLREVAATVSAN 360
Db      301  YNLTSDVGSCTLVCPILHNOEVTAEQTCCKSKCAVCVGLGMEHLREVAATVSAN 360

Qy      301  YNLTSDVGSCTLVCPILHNOEVTAEQTCCKSKCAVCVGLGMEHLREVAATVSAN 360
Db      301  YNLTSDVGSCTLVCPILHNOEVTAEQTCCKSKCAVCVGLGMEHLREVAATVSAN 360

Qy      361  IOEFAGCKKIFGSLAPLPESFDGPASNTAPLOPEOLQVEETLEITGYLYISAMPDLP 420
Db      361  IOEFAGCKKIFGSLAPLPESFDGPASNTAPLOPEOLQVEETLEITGYLYISAMPDLP 420

Qy      361  IOEFAGCKKIFGSLAPLPESFDGPASNTAPLOPEOLQVEETLEITGYLYISAMPDLP 420
Db      361  IOEFAGCKKIFGSLAPLPESFDGPASNTAPLOPEOLQVEETLEITGYLYISAMPDLP 420

Qy      421  DLSVPONLQVIRGRILHNGAVSLTLOGISWGLRSLRELISGLALIHNTHLCPVHTV 480
Db      421  DLSVPONLQVIRGRILHNGAVSLTLOGISWGLRSLRELISGLALIHNTHLCPVHTV 480

Qy      421  DLSVPONLQVIRGRILHNGAVSLTLOGISWGLRSLRELISGLALIHNTHLCPVHTV 480
Db      421  DLSVPONLQVIRGRILHNGAVSLTLOGISWGLRSLRELISGLALIHNTHLCPVHTV 480

Qy      481  PMPOLPENNPOALHTANRPDECEVGBGLACHOUCAHGCMGPRPTCVNCSOFLROEC 540
Db      481  PMPOLPENNPOALHTANRPDECEVGBGLACHOUCAHGCMGPRPTCVNCSOFLROEC 540

Qy      481  PMPOLPENNPOALHTANRPDECEVGBGLACHOUCAHGCMGPRPTCVNCSOFLROEC 540
Db      481  PMPOLPENNPOALHTANRPDECEVGBGLACHOUCAHGCMGPRPTCVNCSOFLROEC 540

Qy      541  VEECRVYQGLPREVYNARHCLPCHPCOPONGSTCFEPADQCVACAHYDPPFCVARC 600
Db      541  VEECRVYQGLPREVYNARHCLPCHPCOPONGSTCFEPADQCVACAHYDPPFCVARC 600

Qy      541  VEECRVYQGLPREVYNARHCLPCHPCOPONGSTCFEPADQCVACAHYDPPFCVARC 600
Db      541  VEECRVYQGLPREVYNARHCLPCHPCOPONGSTCFEPADQCVACAHYDPPFCVARC 600

Qy      601  PSGVKPDLSTYMPKPFDEBGAQPCPINCTHSCVDLDDKCCPAEORASPLTISAVVG 660
Db      601  PSGVKPDLSTYMPKPFDEBGAQPCPINCTHSCVDLDDKCCPAEORASPLTISAVVG 660

Qy      601  PSGVKPDLSTYMPKPFDEBGAQPCPINCTHSCVDLDDKCCPAEORASPLTISAVVG 660
Db      601  PSGVKPDLSTYMPKPFDEBGAQPCPINCTHSCVDLDDKCCPAEORASPLTISAVVG 660

Qy      661  ILVVTGAVRGLIIRKROCKIRKTYMRRLQETELVEPLPSGAMPQAOIRLKEHEL 720
Db      661  ILVVTGAVRGLIIRKROCKIRKTYMRRLQETELVEPLPSGAMPQAOIRLKEHEL 720

Qy      661  ILVVTGAVRGLIIRKROCKIRKTYMRRLQETELVEPLPSGAMPQAOIRLKEHEL 720
Db      661  ILVVTGAVRGLIIRKROCKIRKTYMRRLQETELVEPLPSGAMPQAOIRLKEHEL 720

Qy      721  RKVAVLSGAGTYKGIWIPDGENVKI PVAIKYLRKTSFKANKELIDEAYVNAAGVSP 780
Db      721  RKVAVLSGAGTYKGIWIPDGENVKI PVAIKYLRKTSFKANKELIDEAYVNAAGVSP 780

Qy      721  RKVAVLSGAGTYKGIWIPDGENVKI PVAIKYLRKTSFKANKELIDEAYVNAAGVSP 780
Db      721  RKVAVLSGAGTYKGIWIPDGENVKI PVAIKYLRKTSFKANKELIDEAYVNAAGVSP 780

Qy      781  YVSLGLICLSTVQLVTOIMPYGLLDHVENRGRIGSODLLMCMQIAKMSYLEDVR 840
Db      781  YVSLGLICLSTVQLVTOIMPYGLLDHVENRGRIGSODLLMCMQIAKMSYLEDVR 840
```

```
Db      781  YVSLGLICLSTVQLVTOIMPYGLLDHVENRGRIGSODLLMCMQIAKMSYLEDVR 840
Qy      841  LVHRDLAARNVLYKSPNHVKTIDFGIARLDIDETERYHADGKVPIKMALESILRRFT 900
Db      841  LVHRDLAARNVLYKSPNHVKTIDFGIARLDIDETERYHADGKVPIKMALESILRRFT 900

Qy      901  HOSDWSIGVYTWELMTFGAKYDGI PAREIDPLEKGRLLPOPPCTIDVYIMKCMW 960
Db      901  HOSDWSIGVYTWELMTFGAKYDGI PAREIDPLEKGRLLPOPPCTIDVYIMKCMW 960

Qy      961  INSECRPRRELVSERBARADPORFVVTQNEIDGPASPIDSFTYSLEDDMGDLVDA 1020
Db      961  INSECRPRRELVSERBARADPORFVVTQNEIDGPASPIDSFTYSLEDDMGDLVDA 1020

Qy      1021  EEYLVPOQGFPCDPAPAGAGVHHHRSSSTRSGGDLTLGLEPSEEAAPSPLAPSEG 1080
Db      1021  EEYLVPOQGFPCDPAPAGAGVHHHRSSSTRSGGDLTLGLEPSEEAAPSPLAPSEG 1080

Qy      1081  AGSDVFDGDLGMAKAGLOSPLTHDPSPLQRYSEDPVPLPSETDGVAPLTCSPQREY 1140
Db      1081  AGSDVFDGDLGMAKAGLOSPLTHDPSPLQRYSEDPVPLPSETDGVAPLTCSPQREY 1140

Qy      1141  NQDVAPROPSPREBGLPARAPAGATLBRPKTLSPGKGVVQDVFAFGAVENPEYLPQ 1200
Db      1141  NQDVAPROPSPREBGLPARAPAGATLBRPKTLSPGKGVVQDVFAFGAVENPEYLPQ 1200

Qy      1201  GGAAPQHPPEAPSPAFDNLYYWDODPPERGAPSTFGKPTAENBEYGLDVPV 1255
Db      1201  GGAAPQHPPEAPSPAFDNLYYWDODPPERGAPSTFGKPTAENBEYGLDVPV 1255

RESULT 10
US-10-207-498-6
; Sequence 6, Application US/10207498
; Publication NO. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Kati Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HERGULIN AND HER3
; FILE REFERENCE: 30448:103-US-01
; CURRENT APPLICATION NUMBER: US/10/207, 498
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308, 431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-498-6

Query Match      99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  MELAALCRWGLLALLPFGAASSTOVCTGTMKRLRPSPETHLDMRLHLYOGCOVQGNL 60
Db      1  MELAALCRWGLLALLPFGAASSTOVCTGTMKRLRPSPETHLDMRLHLYOGCOVQGNL 60

Qy      61  ELTYLPTNASLSFLQDIQEVGYVLIHNOVROVPLRLRIVRGTOLEFEDNYALAVLDNG 120
Db      61  ELTYLPTNASLSFLQDIQEVGYVLIHNOVROVPLRLRIVRGTOLEFEDNYALAVLDNG 120

Qy      121  DPLANTPTVPGASPGGRLQLSLRTEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180
Db      121  DPLANTPTVPGASPGGRLQLSLRTEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180

Qy      121  DPLANTPTVPGASPGGRLQLSLRTEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180
Db      121  DPLANTPTVPGASPGGRLQLSLRTEILKGVLIQGNPOLCYODTILMKDIFHKNNOLA 180

Qy      181  LTLIDNRSRAKCPSCMCKGSRCKWSSSDCOSLTTTVACGACGACRCKPLPTDCCHQC 240
Db      181  LTLIDNRSRAKCPSCMCKGSRCKWSSSDCOSLTTTVACGACGACRCKPLPTDCCHQC 240
```

```

Db      181  LTLIDTNRBACHPCSPKCKGRKCBSSBDCQSLTRTVACGACRCKGRLPTDCCHQC 240
Qy      241  AAGCTGPRHSCCLACLFHNSGICELHCPALVTNTDTFESMPNBEGRYTFGASCVTACP 300
Db      241  AAGCTGPRHSCCLACLFHNSGICELHCPALVTNTDTFESMPNBEGRYTFGASCVTACP 300
Qy      301  YNYLSTDVSGCTLVCPHNOEVTADGNORCEKSKPCARCYGIGMHLREVAVTISAN 360
Db      301  YNYLSTDVSGCTLVCPHNOEVTADGNORCEKSKPCARCYGIGMHLREVAVTISAN 360
Qy      361  IOBAGCKKIFGSLAFLPESFDPASNTAFLQEBOLQVFTLEBTGYLISAMPDLP 420
Db      361  IOBAGCKKIFGSLAFLPESFDPASNTAFLQEBOLQVFTLEBTGYLISAMPDLP 420
Qy      421  DLSVFQNLQVIRGRIILHNGAYSLTLQGLISWGLRSLRELSSGLALIHNTHLCPVHTV 480
Db      421  DLSVFQNLQVIRGRIILHNGAYSLTLQGLISWGLRSLRELSSGLALIHNTHLCPVHTV 480
Qy      481  PMDQLFRNHQALHTANRPEDECVGEGLAGHOCACGHCKWGPPTQCVNCSQFLRGQEC 540
Db      481  PMDQLFRNHQALHTANRPEDECVGEGLAGHOCACGHCKWGPPTQCVNCSQFLRGQEC 540
Qy      541  VEECRVLQGLPREYVNAHRLCPHPECOPONGSVTCFGBADQVCAHYNDPPFCVARC 600
Db      541  VEECRVLQGLPREYVNAHRLCPHPECOPONGSVTCFGBADQVCAHYNDPPFCVARC 600
Qy      601  PSQVBDLSYMPIMKFPDEBGAQCPCPINCTHSQVLDKGCRAEQASPLTISVAVVG 660
Db      601  PSQVBDLSYMPIMKFPDEBGAQCPCPINCTHSQVLDKGCRAEQASPLTISVAVVG 660
Qy      661  ILLVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMILKETEL 720
Db      661  ILLVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMILKETEL 720
Qy      721  RKVVLGSGAGFTVYKGIWIPDGENVKIPVALIKVLRNTSPKANKEILDEAYVMAVGSP 780
Db      721  RKVVLGSGAGFTVYKGIWIPDGENVKIPVALIKVLRNTSPKANKEILDEAYVMAVGSP 780
Qy      781  YVSRLLGICLTSTFQVLTQMPYGCCLLDHYRENRGSLGODLNMCMQIAKNGSYLEDVR 840
Db      781  YVSRLLGICLTSTFQVLTQMPYGCCLLDHYRENRGSLGODLNMCMQIAKNGSYLEDVR 840
Qy      841  LVHRDLAARNVLYKSPNHVKITPGLARLLDIDETEHADGKVPKIMALESILRRFT 900
Db      841  LVHRDLAARNVLYKSPNHVKITPGLARLLDIDETEHADGKVPKIMALESILRRFT 900
Qy      901  HOSDWSYGVTVLMTFGAKPYDGIIPAREIPDLLEKGERLPQPTCTIDVYIMKCM 960
Db      901  HOSDWSYGVTVLMTFGAKPYDGIIPAREIPDLLEKGERLPQPTCTIDVYIMKCM 960
Qy      961  IOBECRPRRELVSERBRMARDPQRFVIONEDLGPAASPLDSTFYSLDEDDMGDLVDA 1020
Db      961  IOBECRPRRELVSERBRMARDPQRFVIONEDLGPAASPLDSTFYSLDEDDMGDLVDA 1020
Qy      1021  EETLVQOQGFPCDPAPGAGCMVHHRSSSTSSGGGDLTLGLEPSEBEPASPLAPSEG 1080
Db      1021  EETLVQOQGFPCDPAPGAGCMVHHRSSSTSSGGGDLTLGLEPSEBEPASPLAPSEG 1080
Qy      1081  AASDVPDGLGMAAGLQSLPTHDSPLQRYSEDPTVPLPSETDGVAPLTCSPQREYV 1140
Db      1081  AASDVPDGLGMAAGLQSLPTHDSPLQRYSEDPTVPLPSETDGVAPLTCSPQREYV 1140
Qy      1141  NOPDVAPOPSPREGLPAPARPAATLBRKTLSPKNGVGVDPVAFGAVENPEYLTTPQ 1200
Db      1141  NOPDVAPOPSPREGLPAPARPAATLBRKTLSPKNGVGVDPVAFGAVENPEYLTTPQ 1200
Qy      1201  GGAAPPPHPPAPAPAPVNDLTYWDODPBERGAPSTFGKTPTAENPEYTLGLDVPV 1255
Db      1201  GGAAPPPHPPAPAPAPVNDLTYWDODPBERGAPSTFGKTPTAENPEYTLGLDVPV 1255

```

RESULT 11
US-10-338-730-2

```

; Sequence 2, Application US/10338730
; Publication No. US20030147905A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS
; FILE REFERENCE: 5017C
; CURRENT APPLICATION NUMBER: US/10/338, 730
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: US 09/527,487
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-338-730-2

```

```

Query Match          99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1  METALCRMGLLALLPPGAASQVCTGTDKRLPASPEHLDMLRHLVGGCOVQGNL 60
Db      1  METALCRMGLLALLPPGAASQVCTGTDKRLPASPEHLDMLRHLVGGCOVQGNL 60
Qy      61  EETLVPLNASTLQIOIEVQGVYLAHNOVROVPLORLRIIVGTQLFEDNYALAVDNG 120
Db      61  EETLVPLNASTLQIOIEVQGVYLAHNOVROVPLORLRIIVGTQLFEDNYALAVDNG 120
Qy      121  DPLANTTPYVAGSPGGRLRLSLTEILKGVLIQBNPOLCYODTILMDIIFHKNOILA 180
Db      121  DPLANTTPYVAGSPGGRLRLSLTEILKGVLIQBNPOLCYODTILMDIIFHKNOILA 180
Qy      181  LTLIDTNRBACHPCSPKCKGRKCBSSBDCQSLTRTVACGACRCKGRLPTDCCHQC 240
Db      181  LTLIDTNRBACHPCSPKCKGRKCBSSBDCQSLTRTVACGACRCKGRLPTDCCHQC 240
Qy      241  AAGCTGPRHSCCLACLFHNSGICELHCPALVTNTDTFESMPNBEGRYTFGASCVTACP 300
Db      241  AAGCTGPRHSCCLACLFHNSGICELHCPALVTNTDTFESMPNBEGRYTFGASCVTACP 300
Qy      301  YNYLSTDVSGCTLVCPHNOEVTADGNORCEKSKPCARCYGIGMHLREVAVTISAN 360
Db      301  YNYLSTDVSGCTLVCPHNOEVTADGNORCEKSKPCARCYGIGMHLREVAVTISAN 360
Qy      361  IOBAGCKKIFGSLAFLPESFDPASNTAFLQEBOLQVFTLEBTGYLISAMPDLP 420
Db      361  IOBAGCKKIFGSLAFLPESFDPASNTAFLQEBOLQVFTLEBTGYLISAMPDLP 420
Qy      421  DLSVFQNLQVIRGRIILHNGAYSLTLQGLISWGLRSLRELSSGLALIHNTHLCPVHTV 480
Db      421  DLSVFQNLQVIRGRIILHNGAYSLTLQGLISWGLRSLRELSSGLALIHNTHLCPVHTV 480
Qy      481  PMDQLFRNHQALHTANRPEDECVGEGLAGHOCACGHCKWGPPTQCVNCSQFLRGQEC 540
Db      481  PMDQLFRNHQALHTANRPEDECVGEGLAGHOCACGHCKWGPPTQCVNCSQFLRGQEC 540
Qy      541  VEECRVLQGLPREYVNAHRLCPHPECOPONGSVTCFGBADQVCAHYNDPPFCVARC 600
Db      541  VEECRVLQGLPREYVNAHRLCPHPECOPONGSVTCFGBADQVCAHYNDPPFCVARC 600
Qy      601  PSQVBDLSYMPIMKFPDEBGAQCPCPINCTHSQVLDKGCRAEQASPLTISVAVVG 660
Db      601  PSQVBDLSYMPIMKFPDEBGAQCPCPINCTHSQVLDKGCRAEQASPLTISVAVVG 660
Qy      661  ILLVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMILKETEL 720
Db      661  ILLVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMILKETEL 720
Qy      721  RKVVLGSGAGFTVYKGIWIPDGENVKIPVALIKVLRNTSPKANKEILDEAYVMAVGSP 780

```

Db 721 RKVKVLSGSGFVYKGIWIPDGENVKI PVAIKYLRNTSPKANKETLDEAYVMAGVSP 780
QY 781 YVSRLLGICLTSTVQVLTQMLPQGLDHYRENRGRISODLNMCMQIAKMSYLEDDR 840
Db 781 YVSRLLGICLTSTVQVLTQMLPQGLDHYRENRGRISODLNMCMQIAKMSYLEDDR 840
QY 841 LVHRDLAARVNLVYKSPHVKITDFGLARLLDIDETEVHADGKVPKIMMALESTLRRT 900
Db 841 LVHRDLAARVNLVYKSPHVKITDFGLARLLDIDETEVHADGKVPKIMMALESTLRRT 900
QY 901 HQSDWMSYGTWELMTFGAKPYDGI PARETPDLEKGERLPOPICTIDVYMTVKCM 960
Db 901 HQSDWMSYGTWELMTFGAKPYDGI PARETPDLEKGERLPOPICTIDVYMTVKCM 960
QY 961 IDSECRPRFELVSEFRMARDPQRFVVIQNEBIDGPASPLDSTFYRSLLBEDDMDGLVDA 1020
Db 961 IDSECRPRFELVSEFRMARDPQRFVVIQNEBIDGPASPLDSTFYRSLLBEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCPDPAAGAGMVHRRSSSTRSGGDLTLGLEPSEBEARSPPLAPSEG 1080
Db 1021 EBYLVPOQGFPCPDPAAGAGMVHRRSSSTRSGGDLTLGLEPSEBEARSPPLAPSEG 1080
QY 1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLOKYSBPTVPLPSETDGYVAPLTCSPQPEYV 1140
Db 1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLOKYSBPTVPLPSETDGYVAPLTCSPQPEYV 1140
QY 1141 NOPDVRQPPSPRPGPLPAARPAATLERPKTISPGKGVKDVAFAGAVENPEYLTPO 1200
Db 1141 NOPDVRQPPSPRPGPLPAARPAATLERPKTISPGKGVKDVAFAGAVENPEYLTPO 1200
QY 1201 GGAAPQHPPPAPFPAFDNLVYMDODPPERGAPOSTFGTGTAAENPEYLGIDVPV 1255
Db 1201 GGAAPQHPPPAPFPAFDNLVYMDODPPERGAPOSTFGTGTAAENPEYLGIDVPV 1255

RESULT 12
US-10-322-892-4
Sequence 4, Application US/10322892
Publication No. US20030171257A1
GENERAL INFORMATION:
APPLICANT: STRBL, ROBERT C.
APPLICANT: SNEAD, MALCOLM L.
APPLICANT: XU, JIMMY
APPLICANT: VITETTA, ELLEN S.
APPLICANT: WILK, PETER J.
TITLE OF INVENTION: METHOD AND RELATED COMPOSITION EMPLOYING NANOSTRUCTURES
FILE REFERENCE: M07-505
CURRENT APPLICATION NUMBER: US/10/322,892
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: 60/342,894
PRIOR FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 4
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-10-322-892-4

Query Match 99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELALCGMGLLALLPRGAASOVCTGTDMKRLPASPETHLDMLRLHYGCGVQGNL 60
Db 1 MELALCGMGLLALLPRGAASOVCTGTDMKRLPASPETHLDMLRLHYGCGVQGNL 60
QY 61 ELTYLPTNASLSPLODIOEVGVYLIANNOVROVPLQRLIVRGTOLEEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSPLODIOEVGVYLIANNOVROVPLQRLIVRGTOLEEDNYALAVLDNG 120
QY 121 DPLANTTPTVAGSPGGLRELQSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
Db 121 DPLANTTPTVAGSPGGLRELQSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180

Db 121 DPLANTTPTVAGSPGGLRELQSLTEILKGVLIQRPOLCYODTILMKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMCKGRSCEWBSSEDCQSLTRTVACGACARCKGLPTDCHEQC 240
Db 181 LTLIDNRSRACHPCSPMCKGRSCEWBSSEDCQSLTRTVACGACARCKGLPTDCHEQC 240
QY 241 AAGCTGPKASDCLACHFNHSGTCEHACPALVYNTDTEFSMNPBERYTFGASCYACP 300
Db 241 AAGCTGPKASDCLACHFNHSGTCEHACPALVYNTDTEFSMNPBERYTFGASCYACP 300
QY 301 YNLTSDVGSCTLVCPAHNOEYTABDGTORCEKSKRCACVYGLGMEHLREVAATVSAN 360
Db 301 YNLTSDVGSCTLVCPAHNOEYTABDGTORCEKSKRCACVYGLGMEHLREVAATVSAN 360
QY 361 IQSPACCKIKFGLAFIPESFDGDPASNTAPLOPBOLOVEETLEITGYLISAMPDLP 420
Db 361 IQSPACCKIKFGLAFIPESFDGDPASNTAPLOPBOLOVEETLEITGYLISAMPDLP 420
QY 421 DLSVFQNLQYTRGRILLNCAYSITLQGLISWIGLSRLRELSGLALIHNTHLCPVHTV 480
Db 421 DLSVFQNLQYTRGRILLNCAYSITLQGLISWIGLSRLRELSGLALIHNTHLCPVHTV 480
QY 481 PMOPLRNPQALHTNRNPEDECVBGLACHOLCARGHCKGPRPTOCVNCQFLRGEC 540
Db 481 PMOPLRNPQALHTNRNPEDECVBGLACHOLCARGHCKGPRPTOCVNCQFLRGEC 540
QY 541 VBSGRVLOGLPREYVNAHCLPCHBECOPONGSVTCFGEPAODCVACAHYKDPFCVARC 600
Db 541 VBSGRVLOGLPREYVNAHCLPCHBECOPONGSVTCFGEPAODCVACAHYKDPFCVARC 600
QY 601 PSGVKPDLSTMPYWKPRDEGACQPCINCTHSCVDLDDGCPMEQASPLTISIAYVG 660
Db 601 PSGVKPDLSTMPYWKPRDEGACQPCINCTHSCVDLDDGCPMEQASPLTISIAYVG 660
QY 661 ILLVVLGVYFGLIKRROOKIRKTYRRLLOETELVEPLTPSGAMENOQMRILKETEL 720
Db 661 ILLVVLGVYFGLIKRROOKIRKTYRRLLOETELVEPLTPSGAMENOQMRILKETEL 720
QY 721 RKVKVLSGSGFVYKGIWIPDGENVKI PVAIKYLRNTSPKANKETLDEAYVMAGVSP 780
Db 721 RKVKVLSGSGFVYKGIWIPDGENVKI PVAIKYLRNTSPKANKETLDEAYVMAGVSP 780
QY 781 YVSRLLGICLTSTVQVLTQMLPQGLDHYRENRGRISODLNMCMQIAKMSYLEDDR 840
Db 781 YVSRLLGICLTSTVQVLTQMLPQGLDHYRENRGRISODLNMCMQIAKMSYLEDDR 840
QY 841 LVHRDLAARVNLVYKSPHVKITDFGLARLLDIDETEVHADGKVPKIMMALESTLRRT 900
Db 841 LVHRDLAARVNLVYKSPHVKITDFGLARLLDIDETEVHADGKVPKIMMALESTLRRT 900
QY 901 HQSDWMSYGTWELMTFGAKPYDGI PARETPDLEKGERLPOPICTIDVYMTVKCM 960
Db 901 HQSDWMSYGTWELMTFGAKPYDGI PARETPDLEKGERLPOPICTIDVYMTVKCM 960
QY 961 IDSECRPRFELVSEFRMARDPQRFVVIQNEBIDGPASPLDSTFYRSLLBEDDMDGLVDA 1020
Db 961 IDSECRPRFELVSEFRMARDPQRFVVIQNEBIDGPASPLDSTFYRSLLBEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCPDPAAGAGMVHRRSSSTRSGGDLTLGLEPSEBEARSPPLAPSEG 1080
Db 1021 EBYLVPOQGFPCPDPAAGAGMVHRRSSSTRSGGDLTLGLEPSEBEARSPPLAPSEG 1080
QY 1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLOKYSBPTVPLPSETDGYVAPLTCSPQPEYV 1140
Db 1081 AGSDVDFGDLGMAKAGLQSLPTHDPSPLOKYSBPTVPLPSETDGYVAPLTCSPQPEYV 1140
QY 1141 NOPDVRQPPSPRPGPLPAARPAATLERPKTISPGKGVKDVAFAGAVENPEYLTPO 1200
Db 1141 NOPDVRQPPSPRPGPLPAARPAATLERPKTISPGKGVKDVAFAGAVENPEYLTPO 1200
QY 1201 GGAAPQHPPPAPFPAFDNLVYMDODPPERGAPOSTFGTGTAAENPEYLGIDVPV 1255
Db 1201 GGAAPQHPPPAPFPAFDNLVYMDODPPERGAPOSTFGTGTAAENPEYLGIDVPV 1255

RESULT 13
US-10-272-437A-28
Sequence 28, Application US/10272437A
Publication No. US20030216309A1
GENERAL INFORMATION:
APPLICANT: Krag, David N.
APPLICANT: Petro, Stephanie C.
APPLICANT: Olisio, Lyn
TITLE OF INVENTION: BINDING PEPTIDES SPECIFIC FOR THE EXTRACELLULAR DOMAIN OF ERBB2
TITLE OF INVENTION: US9578987
FILE REFERENCE: V00139, 70056, US
CURRENT APPLICATION NUMBER: US/10/272, 437A
PRIOR FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: 60/329, 183
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patent version 3.1
SEQ ID NO 28
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-10-272-437A-28

Query Match 99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 METALCRMGLLALPPGASTOVCTGTMKRLRASPETHLDMRLHYOGQVVGML 60
DB 1 METALCRMGLLALPPGASTOVCTGTMKRLRASPETHLDMRLHYOGQVVGML 60
QY 61 ELTYLPTNASLFLDIOEVGVYLAHQVQVPLQRLIRYKGTLPEDNTALAVLDNG 120
DB 61 ELTYLPTNASLFLDIOEVGVYLAHQVQVPLQRLIRYKGTLPEDNTALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGVLIQRNPOLCYDITLWKDIFHKNQQA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGVLIQRNPOLCYDITLWKDIFHKNQQA 180
QY 181 LTLIDNRSRACHPCSPMKSGRSGESSEDCSLTRTYCAGGACRCKPLPTDCHBQC 240
DB 181 LTLIDNRSRACHPCSPMKSGRSGESSEDCSLTRTYCAGGACRCKPLPTDCHBQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNDTFESMNPREGRTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNDTFESMNPREGRTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPLNQVETABDGTQRCCKSPCARVCYGLAMEHLREVRATVSAN 360
DB 301 YNYLSTDVGSCTLVCPLNQVETABDGTQRCCKSPCARVCYGLAMEHLREVRATVSAN 360
QY 361 IOEPAGCKKIRGSLAFLESPFGDPASNTAPLQPELOVFEETLEITGLYISAMPDILP 420
DB 361 IOEPAGCKKIRGSLAFLESPFGDPASNTAPLQPELOVFEETLEITGLYISAMPDILP 420
QY 421 DLSVFNQVLRGRILHNGAVSLTLOGISWLGSLBELGSSGALITHNHILCFYTHV 480
DB 421 DLSVFNQVLRGRILHNGAVSLTLOGISWLGSLBELGSSGALITHNHILCFYTHV 480
QY 481 PMDOLFRNPHOALHTANRPEDECVGEGSLAQHQLCARGHCMGPGFTQCVNCSQPLRGDEC 540
DB 481 PMDOLFRNPHOALHTANRPEDECVGEGSLAQHQLCARGHCMGPGFTQCVNCSQPLRGDEC 540
QY 541 VEECKVILQGLPREYVNAHCLPCHRECOFONGSVTCFGEADQCVACHYKDPFCVARC 600
DB 541 VEECKVILQGLPREYVNAHCLPCHRECOFONGSVTCFGEADQCVACHYKDPFCVARC 600
QY 601 PSGVPELSTYMPIMKFPDEBEGACOPCPINCTHSQVLDLDDKGPAPORASPLTSISAVVG 660
DB 601 PSGVPELSTYMPIMKFPDEBEGACOPCPINCTHSQVLDLDDKGPAPORASPLTSISAVVG 660

QY 661 ILLVVLGVVFGILIRROQKIRKTYMRLLQSTELVEPLTPSGAMPNOAMRLKETEL 720
DB 661 ILLVVLGVVFGILIRROQKIRKTYMRLLQSTELVEPLTPSGAMPNOAMRLKETEL 720
QY 721 RKYKVLGSAFGVYVGIWIPDEENKIPALIVALENTPKANKELIDEAYVMAVGSP 780
DB 721 RKYKVLGSAFGVYVGIWIPDEENKIPALIVALENTPKANKELIDEAYVMAVGSP 780
QY 781 YSRLILGICLTSTVQVLTQMLPQGLLDHVBENRGLSGODLNMCMQIAKMSYLEDVR 840
DB 781 YSRLILGICLTSTVQVLTQMLPQGLLDHVBENRGLSGODLNMCMQIAKMSYLEDVR 840
QY 841 LVHRDLAANVILKSNHYKITDFGLARLLDIDETEVHADGQVPIKMALESILRRFT 900
DB 841 LVHRDLAANVILKSNHYKITDFGLARLLDIDETEVHADGQVPIKMALESILRRFT 900
QY 901 HQSDWVSXVATWELMTFCAKPYDGIPIAREIPDLKGERLPQPICTIDVYIMVXCM 960
DB 901 HQSDWVSXVATWELMTFCAKPYDGIPIAREIPDLKGERLPQPICTIDVYIMVXCM 960
QY 961 IDSECRPRELVSERSRMAPDPORFVIONEDLGPASPLDSTFFYSLLEDMDGLVDA 1020
DB 961 IDSECRPRELVSERSRMAPDPORFVIONEDLGPASPLDSTFFYSLLEDMDGLVDA 1020
QY 1021 EBYLVPOQGFCCPDPAAGAVHHRHRSSTRSGGDLTLGLEPSBEAPRSPAPSEG 1080
DB 1021 EBYLVPOQGFCCPDPAAGAVHHRHRSSTRSGGDLTLGLEPSBEAPRSPAPSEG 1080
QY 1081 AGSDVFDGDLGMAKGLQSLPTHDBPLQRYSEDPVLPSETDGVVAPLTCSPOREYV 1140
DB 1081 AGSDVFDGDLGMAKGLQSLPTHDBPLQRYSEDPVLPSETDGVVAPLTCSPOREYV 1140
QY 1141 NQPDVAPQPSRREGPLPARBAGATLBRPKLSPEKSNVVDVAFGAVENPEYVTPQ 1200
DB 1141 NQPDVAPQPSRREGPLPARBAGATLBRPKLSPEKSNVVDVAFGAVENPEYVTPQ 1200
QY 1201 GGAAPQHPHPPAPSAFDNLTYWDDPPERGAPESTFKTPTAENPEYGLDVPV 1255
DB 1201 GGAAPQHPHPPAPSAFDNLTYWDDPPERGAPESTFKTPTAENPEYGLDVPV 1255

RESULT 14
US-10-117-937-594
Sequence 594, Application US/10117937
Publication No. US20030220239A1
GENERAL INFORMATION:
APPLICANT: CTL IMMO NO THERAPIES CORP.
APPLICANT: SIMARD, John, J.L.
APPLICANT: DIAMOND, David, C.
APPLICANT: Liu, Liding
APPLICANT: Xie, Zhidong
TITLE OF INVENTION: EPTOPE SEQUENCES
FILE REFERENCE: CTIIMM.027A
CURRENT APPLICATION NUMBER: US/10/117, 937
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/282, 211
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/337, 017
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/363, 210
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 602
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 594
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-10-117-937-594

Query Match 99.9%; Score 6806; DB 12; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPFGAASVQVCTGDMKRLPASPETHDMLRHLVYQGCQVQGNL 60
 DB 1 MELAALCRWGLLALLPFGAASVQVCTGDMKRLPASPETHDMLRHLVYQGCQVQGNL 60
 QY 61 ELTYLPTNASTSLFODIOEVQGYVLIANQVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120
 DB 61 ELTYLPTNASTSLFODIOEVQGYVLIANQVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120
 QY 121 DPLANTTPVTGASPGGLRELQLSLREILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
 DB 121 DPLANTTPVTGASPGGLRELQLSLREILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
 QY 181 LTLIDNRSRACHPCSPMCKGSRCKWGBSSBDQSLTRTVACGACRCKGPLPTDCCHBQC 240
 DB 181 LTLIDNRSRACHPCSPMCKGSRCKWGBSSBDQSLTRTVACGACRCKGPLPTDCCHBQC 240
 QY 241 MAGCTGRKSHDCLACHFNHSGICELHCPALVYNTDTPFESMPNBERYTFGASCVTACP 300
 DB 241 MAGCTGRKSHDCLACHFNHSGICELHCPALVYNTDTPFESMPNBERYTFGASCVTACP 300
 QY 301 YNLTSTDVGSCTVCPPLHNOEVTABDGTORCEKSKFCARVCYGLGMEHLREYAVTSAN 360
 DB 301 YNLTSTDVGSCTVCPPLHNOEVTABDGTORCEKSKFCARVCYGLGMEHLREYAVTSAN 360
 QY 361 IQBFAGCKKIFGSLAFIPESFDDGPASNTAPLOEQLQVETLEITGYLIYSAMPDSL 420
 DB 361 IQBFAGCKKIFGSLAFIPESFDDGPASNTAPLOEQLQVETLEITGYLIYSAMPDSL 420
 QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRSLREISGLALIHNTHLCFVHTV 480
 DB 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRSLREISGLALIHNTHLCFVHTV 480
 QY 481 PMDOLFPNPHQALLHTANRPEDECVGEGGLACHOLCARGHCWGPPTQCVNCSQPLRQEC 540
 DB 481 PMDOLFPNPHQALLHTANRPEDECVGEGGLACHOLCARGHCWGPPTQCVNCSQPLRQEC 540
 QY 541 VEECRVLQGLPREVYNAHCLPCHBECQOPONGSVTCFGEADQCVACAHYKDPFCVAKC 600
 DB 541 VEECRVLQGLPREVYNAHCLPCHBECQOPONGSVTCFGEADQCVACAHYKDPFCVAKC 600
 QY 601 PEGVYKPLSTMPYKPFDEBGAQCPICNTHSCVDLDKGCRAEORASPLTSTISAVG 660
 DB 601 PEGVYKPLSTMPYKPFDEBGAQCPICNTHSCVDLDKGCRAEORASPLTSTISAVG 660
 QY 661 ILLVYLVGVVFGILIKRQOKIKRYTMRRLQETELVEPLTPSGAMPNOAMRILKXTEL 720
 DB 661 ILLVYLVGVVFGILIKRQOKIKRYTMRRLQETELVEPLTPSGAMPNOAMRILKXTEL 720
 QY 721 RKYKVLGSGAFGTVYKGIWIPDGENVKI PVAIKVLRENTSFKANKEILDEAYVAGVSP 780
 DB 721 RKYKVLGSGAFGTVYKGIWIPDGENVKI PVAIKVLRENTSFKANKEILDEAYVAGVSP 780
 QY 781 YVSRLLGICLTSTQVLTQMLPYGCLLDHRENRGRGLSODLWMCQIAGMSYLIBDVR 840
 DB 781 YVSRLLGICLTSTQVLTQMLPYGCLLDHRENRGRGLSODLWMCQIAGMSYLIBDVR 840
 QY 841 LVHRDLAABVNLVYSPHNVKLTDFGLARLLDIDETBYHAGDGKYPIMMALESILRRFT 900
 DB 841 LVHRDLAABVNLVYSPHNVKLTDFGLARLLDIDETBYHAGDGKYPIMMALESILRRFT 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLKGERLPOPPICITIDVYMIWVCKM 960
 DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLKGERLPOPPICITIDVYMIWVCKM 960
 QY 961 IDSECRPFRRELVEBFSHMAADPOAFVYIQNEBDLGPASPLDSTFRSLLEDDMDGLVDA 1020
 DB 961 IDSECRPFRRELVEBFSHMAADPOAFVYIQNEBDLGPASPLDSTFRSLLEDDMDGLVDA 1020
 QY 1021 BEYLVPOGFPFCPPDAPAGGAVHHRHSSTRSGGDLTGLBESSEBAPRSLAPBSG 1080
 DB 1021 BEYLVPOGFPFCPPDAPAGGAVHHRHSSTRSGGDLTGLBESSEBAPRSLAPBSG 1080
 QY 1081 AGSDVFPDGLMGAAKGLQSLPTHDPSPLOQRYSDPTVLPSETDGYVAPLTCSPQPEYV 1140

DB 1081 AGSDVFPDGLMGAAKGLQSLPTHDPSPLOQRYSDPTVLPSETDGYVAPLTCSPQPEYV 1140
 QY 1141 NOPDVAPQPPSPREGELPAARAGATLERPKTLPKNGVYKDVAFGAVENPEYLTPO 1200
 DB 1141 NOPDVAPQPPSPREGELPAARAGATLERPKTLPKNGVYKDVAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPHPPAPSPAFDNLVYWDODPPERGAPSTFKGTPTAENPEYLGLDVAV 1255
 DB 1201 GGAAPQHPHPPAPSPAFDNLVYWDODPPERGAPSTFKGTPTAENPEYLGLDVAV 1255
 RESULT 15
 US-10-435-696-36
 ; Sequence 36, Application US/10435696
 ; Publication No. US20040018525A1
 GENERAL INFORMATION:
 ; APPLICANT: Wirtz, Ralph
 ; APPLICANT: Munnes, Marc
 ; APPLICANT: Kallabie, Harald
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS
 ; TITLE OF INVENTION: PREVENTION AND TREATMENT OF MALIGNANT NEOPLASIA
 ; FILE REFERENCE: Lea 36 108
 ; CURRENT APPLICATION NUMBER: US/10/435,696
 ; CURRENT FILING DATE: 2003-05-09
 ; PRIOR APPLICATION NUMBER: EP03003112.4
 ; PRIOR FILING DATE: 2003-02-13
 ; PRIOR APPLICATION NUMBER: EP02010291.9
 ; PRIOR FILING DATE: 2002-05-21
 ; NUMBER OF SEQ ID NOS: 314
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 36
 ; LENGTH: 1255
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-435-696-36
 Query Match 99.9%; Score 6806; DB 12; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MELAALCRWGLLALLPFGAASVQVCTGDMKRLPASPETHDMLRHLVYQGCQVQGNL 60
 DB 1 MELAALCRWGLLALLPFGAASVQVCTGDMKRLPASPETHDMLRHLVYQGCQVQGNL 60
 QY 61 ELTYLPTNASTSLFODIOEVQGYVLIANQVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120
 DB 61 ELTYLPTNASTSLFODIOEVQGYVLIANQVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120
 QY 121 DPLANTTPVTGASPGGLRELQLSLREILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
 DB 121 DPLANTTPVTGASPGGLRELQLSLREILKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
 QY 181 LTLIDNRSRACHPCSPMCKGSRCKWGBSSBDQSLTRTVACGACRCKGPLPTDCCHBQC 240
 DB 181 LTLIDNRSRACHPCSPMCKGSRCKWGBSSBDQSLTRTVACGACRCKGPLPTDCCHBQC 240
 QY 241 MAGCTGRKSHDCLACHFNHSGICELHCPALVYNTDTPFESMPNBERYTFGASCVTACP 300
 DB 241 MAGCTGRKSHDCLACHFNHSGICELHCPALVYNTDTPFESMPNBERYTFGASCVTACP 300
 QY 301 YNLTSTDVGSCTVCPPLHNOEVTABDGTORCEKSKFCARVCYGLGMEHLREYAVTSAN 360
 DB 301 YNLTSTDVGSCTVCPPLHNOEVTABDGTORCEKSKFCARVCYGLGMEHLREYAVTSAN 360
 QY 361 IQBFAGCKKIFGSLAFIPESFDDGPASNTAPLOEQLQVETLEITGYLIYSAMPDSL 420
 DB 361 IQBFAGCKKIFGSLAFIPESFDDGPASNTAPLOEQLQVETLEITGYLIYSAMPDSL 420
 QY 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRSLREISGLALIHNTHLCFVHTV 480
 DB 421 DLSVFQNLQVIRGRILHNGAYSLTLOGLISWGLRSLREISGLALIHNTHLCFVHTV 480

QY 481 PWDOLFRNPHOALLHTANRPEDECVGEGLACHOLCARGHCMGPRPTQCVNCSQFLRGOC 540
DB 481 PWDOLFRNPHOALLHTANRPEDECVGEGLACHOLCARGHCMGPRPTQCVNCSQFLRGOC 540
QY 541 VEECVTLQGLPRBYVVARHCLPCHPECOPONGSVTCFEPADOCVACAHYKDPFCVAC 600
DB 541 VEECVTLQGLPRBYVVARHCLPCHPECOPONGSVTCFEPADOCVACAHYKDPFCVAC 600
QY 601 PSGVXPDLSPYMPIMKPEDEEGACQPCPINCTHSCVDLDDKGCRAEORASPLTSIIISAVVG 660
DB 601 PSGVXPDLSPYMPIMKPEDEEGACQPCPINCTHSCVDLDDKGCRAEORASPLTSIIISAVVG 660
QY 661 ILLVVVLGVVFGILLIKRROOKI RKTMRRLLOETELVEPLTPSGAMPNOQRILKETEL 720
DB 661 ILLVVVLGVVFGILLIKRROOKI RKTMRRLLOETELVEPLTPSGAMPNOQRILKETEL 720
QY 721 RKYKVLGSGAFSTVYKGIWIPGENVKIPVAIKVLEBNTSPKANKIILDEAYMAGVGP 780
DB 721 RKYKVLGSGAFSTVYKGIWIPGENVKIPVAIKVLEBNTSPKANKIILDEAYMAGVGP 780
QY 781 YVSRLLGICLTSTVOQLVTOJMPYGLLDHVRENRGLSGODLLNMCQIAKMSYLEDV 840
DB 781 YVSRLLGICLTSTVOQLVTOJMPYGLLDHVRENRGLSGODLLNMCQIAKMSYLEDV 840
QY 841 LVHRDLAARNVLYKSPNHKLTDFGLARLLDIDETRYHADGKVPKIMMALISILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHKLTDFGLARLLDIDETRYHADGKVPKIMMALISILRRFT 900
QY 901 HOSDVMSYGVTVWMLTFGAKPYDGI BAREIPDLKEXERLPOPICTIDVYIMVCMW 960
DB 901 HOSDVMSYGVTVWMLTFGAKPYDGI BAREIPDLKEXERLPOPICTIDVYIMVCMW 960
QY 961 IDSECRPRRELVSERSMARDPORFVVIQNEBLAGPASPILDSTFYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRRELVSERSMARDPORFVVIQNEBLAGPASPILDSTFYRSLLEDDMDGLVDA 1020
QY 1021 BEYLVPOCGFFCPDPAFGAGVHHHRSSSTRSGGDLTLGLEPSEBEAPRSPPLASEG 1080
DB 1021 BEYLVPOCGFFCPDPAFGAGVHHHRSSSTRSGGDLTLGLEPSEBEAPRSPPLASEG 1080
QY 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSEDPVPLPSBTGCVAPLTCSPQPEV 1140
DB 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSEDPVPLPSBTGCVAPLTCSPQPEV 1140
QY 1141 NOPDVRPOPSPREBEPPLAARPAATLERPKTILSPKNGVVKOVFAFGAVENPEYLTPQ 1200
DB 1141 NOPDVRPOPSPREBEPPLAARPAATLERPKTILSPKNGVVKOVFAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQPHPPPAFPAFDNLVYWDQDPPRGAPESTFKGPTAENPEYLGLDVVP 1255
DB 1201 GGAAPQPHPPPAFPAFDNLVYWDQDPPRGAPESTFKGPTAENPEYLGLDVVP 1255

Search completed: February 9, 2004, 16:22:57
Job time : 50 secs

XX MPI; 1996-455361/45.
 DR N-PSDB; AAT40739.
 XX DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
 PT treatment of malignancies with which the HER-2/neu oncogene is
 PT associated.
 XX
 PS Claim 2; Page 56-61; 71pp; English.
 XX
 CC Human HER-2/neu protein (AAM01111), also called p185 or c-erbB2, is
 CC the product of the HER-2/neu oncogene (see also AAT40739). The
 CC protein is over-expressed in various cancers, including breast,
 CC ovarian, colon, lung and prostate. The intracellular domain of the
 CC protein can be used to immunise an animal against a malignancy with
 CC which the oncogene is associated. The polypeptide can be produced
 CC in transformed host cells for use in immunisation. Alternatively,
 CC animal cells are transfected in vivo or ex vivo with a viral vector
 CC that directs expression of the polypeptide.
 CC
 SQ Sequence 1255 AA;
 Query Match 100.0%; Score 6815; DB 17; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELAALCRNGLLALALPGAASTQVCTGDMKRLPASPETHLDMRLHYQGQVQGNL 60
 DB 1 MELAALCRNGLLALALPGAASTQVCTGDMKRLPASPETHLDMRLHYQGQVQGNL 60
 QY 61 ELTYLPTNALSLSLQDIQEVQYVLAHNOVQVPLQRLRIVRGTQLFEDNALAVLDNG 120
 DB 61 ELTYLPTNALSLSLQDIQEVQYVLAHNOVQVPLQRLRIVRGTQLFEDNALAVLDNG 120
 QY 121 DPLNNTPTVTSAPGRLRELQRLSTELIKGVLIQNPQLCYODTILMDIFHKNNQLA 180
 DB 121 DPLNNTPTVTSAPGRLRELQRLSTELIKGVLIQNPQLCYODTILMDIFHKNNQLA 180
 QY 181 LTLIDNRSRACHPCGSPMCKSGRCWGSSESDCSLFTVQAGGACARCKGPLPTDCHEOC 240
 DB 181 LTLIDNRSRACHPCGSPMCKSGRCWGSSESDCSLFTVQAGGACARCKGPLPTDCHEOC 240
 QY 241 AAGCTGPKISDCLACHFNHSGICEHLCPALVTYNTDFESMNPBGRYTFGASCYTAAP 300
 DB 241 AAGCTGPKISDCLACHFNHSGICEHLCPALVTYNTDFESMNPBGRYTFGASCYTAAP 300
 QY 301 YNYLSTDVSGCTVCPHNOEYTABDGTORCKSCSPCAVCYGLGMEHLREYRAVTSAN 360
 DB 301 YNYLSTDVSGCTVCPHNOEYTABDGTORCKSCSPCAVCYGLGMEHLREYRAVTSAN 360
 QY 361 IOEFACCKTIFGSLAFLPESPDGDPASNTAPLOPBOLOVETLEITGYLYISAMPDLP 420
 DB 361 IOEFACCKTIFGSLAFLPESPDGDPASNTAPLOPBOLOVETLEITGYLYISAMPDLP 420
 QY 421 DLSVFONLOVIRGRIIHNNGAYSLTLOGLISWLGSLSLRELSSGLALIHNTHLCPVHTV 480
 DB 421 DLSVFONLOVIRGRIIHNNGAYSLTLOGLISWLGSLSLRELSSGLALIHNTHLCPVHTV 480
 QY 481 PMDOLFRNPHQALLHNRANREDECVGEGLAHQOLCARHGWGSPGTCVNCSPFLGQEC 540
 DB 481 PMDOLFRNPHQALLHNRANREDECVGEGLAHQOLCARHGWGSPGTCVNCSPFLGQEC 540
 QY 541 VEECRVLOGLPREYVNAHCLPCGPECQPNQSVTCGFPADQCVACAHXDPFVYARC 600
 DB 541 VEECRVLOGLPREYVNAHCLPCGPECQPNQSVTCGFPADQCVACAHXDPFVYARC 600
 QY 601 PSGVKDLSYMPYWKPEDEGACQPCINCTHSCVDLDKGCAPABQASPLTYSIIISAVG 660
 DB 601 PSGVKDLSYMPYWKPEDEGACQPCINCTHSCVDLDKGCAPABQASPLTYSIIISAVG 660
 QY 661 ILLVVLGVVFGILIRROOKIKRTYMRRLLOSTELVEPLTSGGAMPNOQRIKXETEL 720
 DB 661 ILLVVLGVVFGILIRROOKIKRTYMRRLLOSTELVEPLTSGGAMPNOQRIKXETEL 720

QY 721 RAYKVLGSGAFGTVYKGIWIPDGENYKIPVAIKVLEBNTSPKANKELDEAYVNAVGSP 780
 DB 721 RAYKVLGSGAFGTVYKGIWIPDGENYKIPVAIKVLEBNTSPKANKELDEAYVNAVGSP 780
 QY 781 YNSRLIGICTSTVQVLTQMLPFGCLLDHYRENRGLGSDLLNWCQIAKGSYLEDYR 840
 DB 781 YNSRLIGICTSTVQVLTQMLPFGCLLDHYRENRGLGSDLLNWCQIAKGSYLEDYR 840
 QY 841 LVHRDLAARNVLYKSNHYKITDPGLARLLDIDETRHADGGKVPIMKMLLESTLRRFT 900
 DB 841 LVHRDLAARNVLYKSNHYKITDPGLARLLDIDETRHADGGKVPIMKMLLESTLRRFT 900
 QY 901 HOSDVSXGVTWELMTFGAKPYDGIIPAREIPDLKEGERTPOPICTIDVYIMVYCMW 960
 DB 901 HOSDVSXGVTWELMTFGAKPYDGIIPAREIPDLKEGERTPOPICTIDVYIMVYCMW 960
 QY 961 IDSECRPRRELVSERSMARDPQRFVYIQNEDLGPASPLDSTFYSLLEDDMDGLVDA 1020
 DB 961 IDSECRPRRELVSERSMARDPQRFVYIQNEDLGPASPLDSTFYSLLEDDMDGLVDA 1020
 QY 1021 EBYLVPQGFPCDPAPAGAGVHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
 DB 1021 EBYLVPQGFPCDPAPAGAGVHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
 QY 1081 AGSDVFDGDLGMAAGLQSLPTHDPSPLORYSEDEPTVPLPSETDGYVAPLTCSPQREYV 1140
 DB 1081 AGSDVFDGDLGMAAGLQSLPTHDPSPLORYSEDEPTVPLPSETDGYVAPLTCSPQREYV 1140
 QY 1141 NQPDVAPQPPSPREGPLPAARPAATLERPKTISPGKNGVYKOVFAFGAVENPEYLTPO 1200
 DB 1141 NQPDVAPQPPSPREGPLPAARPAATLERPKTISPGKNGVYKOVFAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPHPAPSPAFNDLYYWDODPPERGAPSTFKGPTAENPEYLGDLVPV 1255
 DB 1201 GGAAPQHPHPAPSPAFNDLYYWDODPPERGAPSTFKGPTAENPEYLGDLVPV 1255

RESULT 2
 ID AAM92406 standard; Protein; 1255 AA.
 XX AAM92406;
 AC AAM92406;
 DT 21-APR-1999 (first entry)
 XX
 DE Human HER-2/neu oncogene protein.
 XX
 KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
 KW malignancy; treatment; tumour.
 OS Homo sapiens.
 FH
 FT Key
 FT Region Location/Qualifiers
 FT 676..1255
 FT /note="region which elicits immune response"
 PN US5869445-A.
 XX
 PD 09-FEB-1999.
 XX
 PF 01-APR-1996; 96US-0625101.
 XX
 PR 01-APR-1996; 96US-0625101.
 PR 17-MAR-1993; 93US-0033644.
 PR 12-AUG-1993; 93US-0106112.
 PR 31-MAR-1995; 95US-0414417.
 XX
 PA (UNIM) UNIM WASHINGTON.
 XX
 PI Cheever MA, Distie ML;
 XX MPI; 1999-152835/13.

DR N-PSDB; AAX01912.

XX Use of HER-2/neu polypeptides - for eliciting an immune response to
PT an HER-2/neu associated malignancy, particularly for treating or
PT preventing tumours

PS Claim 3; Column 31-38; 26pp; English.

CC This sequence represents the human HER-2/neu oncogene protein. A fragment
CC of this protein is used in a method for eliciting or enhancing an immune
CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC B cells to produce an immune response to the HER-2/neu protein. The
CC method can be used for immunisation against a malignancy in which the
CC HER-2/neu oncogene is associated and in the treatment of an existing
CC tumour, or to prevent tumour occurrence or recurrence.

CC Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 20; Length 1255;
Match Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 METALALCKRWGLLALALPAGASTVCTGTDMLRLPASPETHLDMRLHYOGQVYQGNL 60
DB 1 METALALCKRWGLLALALPAGASTVCTGTDMLRLPASPETHLDMRLHYOGQVYQGNL 60
QY 61 ELTYLPTNASLFLDIOEVQGVYLAHNQVRVPLQRLIRVGTQLFEDNTVALVLDNG 120
DB 61 ELTYLPTNASLFLDIOEVQGVYLAHNQVRVPLQRLIRVGTQLFEDNTVALVLDNG 120
QY 121 DPLNNTPTPTGASPGSLRELQRLSLTEILKGVYLRNPNQLCYQDTILWKDIFHKNOQA 180
DB 121 DPLNNTPTPTGASPGSLRELQRLSLTEILKGVYLRNPNQLCYQDTILWKDIFHKNOQA 180
QY 121 DPLNNTPTPTGASPGSLRELQRLSLTEILKGVYLRNPNQLCYQDTILWKDIFHKNOQA 180
DB 121 DPLNNTPTPTGASPGSLRELQRLSLTEILKGVYLRNPNQLCYQDTILWKDIFHKNOQA 180
QY 181 LTLIDTNSRAHCPSPCKSGRCWGSSESDCOSLTRVCAGGCAKCKGPLEPTDCHEOC 240
DB 181 LTLIDTNSRAHCPSPCKSGRCWGSSESDCOSLTRVCAGGCAKCKGPLEPTDCHEOC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTYTFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTYTFGASCYTACP 300
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTYTFGASCYTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPREGRTYTFGASCYTACP 300
QY 301 YNYLSTDVGSCTLVCPHNOVYTABDGTORCEKSCPCARVCYGLAMEHLREYAVTSAN 360
DB 301 YNYLSTDVGSCTLVCPHNOVYTABDGTORCEKSCPCARVCYGLAMEHLREYAVTSAN 360
QY 361 IOBPAGCKKIFGSLAFLPESPDGDPASNTAPLOPBLQVFTLBEITVGLYSAMPDSLIP 420
DB 361 IOBPAGCKKIFGSLAFLPESPDGDPASNTAPLOPBLQVFTLBEITVGLYSAMPDSLIP 420
QY 421 DLSVFNQTLQVIRGRILHNGAYSLTLQGLISWLGRLSLBELSGSLALHHNTHLCPVHTV 480
DB 421 DLSVFNQTLQVIRGRILHNGAYSLTLQGLISWLGRLSLBELSGSLALHHNTHLCPVHTV 480
QY 481 FMDQLFRNPHQALHTANRPEDECVGEGLAGHQLCARHGWPGPTOCVNCQSLRGDQC 540
DB 481 FMDQLFRNPHQALHTANRPEDECVGEGLAGHQLCARHGWPGPTOCVNCQSLRGDQC 540
QY 541 VEECRVLOGLPREYVNAHCLFCHPECOFQNGSVTCFGBADQCVAHAYKDPFVAVAC 600
DB 541 VEECRVLOGLPREYVNAHCLFCHPECOFQNGSVTCFGBADQCVAHAYKDPFVAVAC 600
QY 601 PSGVPRDLSYMPIMKPPDEBACOPCPICHTSCVDLDDKGPABORASPLTISIISAVVG 660
DB 601 PSGVPRDLSYMPIMKPPDEBACOPCPICHTSCVDLDDKGPABORASPLTISIISAVVG 660
QY 661 ILVVVLGVVFGLILKRRQOKIRKTYMRLLQETELVEPLTSGAMPNPAQRIILKETEL 720
DB 661 ILVVVLGVVFGLILKRRQOKIRKTYMRLLQETELVEPLTSGAMPNPAQRIILKETEL 720
QY 721 RKVKTLGSGARCTYVKGIMIPEGENVKIIVALKVLRNTSPKANGIILDEAVYMAVGSP 780
DB 721 RKVKTLGSGARCTYVKGIMIPEGENVKIIVALKVLRNTSPKANGIILDEAVYMAVGSP 780

```

```

QY 781 YVSRLLIGICTSTVQVLTQMLPVGCLLDHYRENRGLSGQDLINMCQIAKNGSYLEDVR 840
DB 781 YVSRLLIGICTSTVQVLTQMLPVGCLLDHYRENRGLSGQDLINMCQIAKNGSYLEDVR 840
QY 841 LVHRDLAANVTVKSNHVKITDFGLARLLDIDETEVHADGCVPIKMALESILRRFT 900
DB 841 LVHRDLAANVTVKSNHVKITDFGLARLLDIDETEVHADGCVPIKMALESILRRFT 900
QY 901 HOSDWSYGVTWELMTFGAKPYDGI PAREIPDLKEGRLPOPICTIDVYIMVCMW 960
DB 901 HOSDWSYGVTWELMTFGAKPYDGI PAREIPDLKEGRLPOPICTIDVYIMVCMW 960
QY 961 ISECRPREPRLVSESRNARDPQRFVLIQNEDLGPASPLDSTFYSLLEDMDMDGLVDA 1020
DB 961 ISECRPREPRLVSESRNARDPQRFVLIQNEDLGPASPLDSTFYSLLEDMDMDGLVDA 1020
QY 1021 ERYLVPQGFPCPDPAFGAGVHHRRSSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
DB 1021 ERYLVPQGFPCPDPAFGAGVHHRRSSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
QY 1081 AGSDVFDGDLGMAKGLQSLPTHDSPLQRYSEDPVLPSEETDGVAPLTCSPQPEYV 1140
DB 1081 AGSDVFDGDLGMAKGLQSLPTHDSPLQRYSEDPVLPSEETDGVAPLTCSPQPEYV 1140
QY 1141 NQPDVAPQPPSPREBGLPAARPAATLERPKTLSPKNGVYKDVFAFGAVENPEYLTPQ 1200
DB 1141 NQPDVAPQPPSPREBGLPAARPAATLERPKTLSPKNGVYKDVFAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPHPAPSPAFDNLTYWDODPPERGAPSPSTFKGTPTAENPEYGLDVPV 1255
DB 1201 GGAAPQHPHPAPSPAFDNLTYWDODPPERGAPSPSTFKGTPTAENPEYGLDVPV 1255

```

RESULT 3
AAB21198
ID AAB21198 standard; protein; 1255 AA.
AC AAB21198;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu protein.
XX
KW Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
XX
OS Homo sapiens.
XX
PN MO200044899-A1.
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000MO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
XX
PA (CORI-) CORIYA CORP.
PA (SMIK-) SMITHLINE BEECHAM.
PI Cheever MA, Gheysen D;
XX
DR WPI: 2000-505976/45.
DR N-PSDB; AAB89736.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
XX prostate cancers -
XX
PS Claim 52; Fig 7; 128pp; English.
CC The present sequence is the human HER-2/neu protein. It is a member of

CC the tyrosine kinase family of receptor-like glycoproteins and shows
CC homology to the epidermal growth factor receptor (EGFR). It probably
CC plays a part in cell growth and/or differentiation. The HER-2/neu
CC gene is an oncogene. An HER-2/neu fusion protein comprising a
CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
CC
XX

Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 21; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METAAACRMGLLALLPFGAASSTOYCTGTDMKRLPASPETHLDMLHLYOGCQVVOGNTL 60
DB 1 METAAACRMGLLALLPFGAASSTOYCTGTDMKRLPASPETHLDMLHLYOGCQVVOGNTL 60
QY 61 ELTYLPTNASTLSPFADIIQEVGYVLIAMNOVROVPLQRLRIVGTOLFEDNYVALAVDNG 120
DB 61 ELTYLPTNASTLSPFADIIQEVGYVLIAMNOVROVPLQRLRIVGTOLFEDNYVALAVDNG 120
QY 121 DPLNNTPTVYGASPGSLRELQRLSLTEILKGVLLIQNPOLCYODITLMDIPIKNNQLA 180
DB 121 DPLNNTPTVYGASPGSLRELQRLSLTEILKGVLLIQNPOLCYODITLMDIPIKNNQLA 180
QY 181 LTIIDNRSAPACIPSCPMCKGSRGWSBSBDCOSLITTVCAAGCARCKGLPTDCHEOC 240
DB 181 LTIIDNRSAPACIPSCPMCKGSRGWSBSBDCOSLITTVCAAGCARCKGLPTDCHEOC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICELCPALVTNTDTPSPMPNBSGRYTFGASCYACR 300
DB 241 AAGCTGPKHSDCLACHFNHSGICELCPALVTNTDTPSPMPNBSGRYTFGASCYACR 300
QY 301 YNTLSTDVSGSCTLVCPHNOEVTAEQGTORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
DB 301 YNTLSTDVSGSCTLVCPHNOEVTAEQGTORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
QY 361 IOEPFAGCKKIFGSLAPFESFDDPANTAPLOPBOLOVETLEETLIGLYISAMPDLP 420
DB 361 IOEPFAGCKKIFGSLAPFESFDDPANTAPLOPBOLOVETLEETLIGLYISAMPDLP 420
QY 421 DLSVFQRLQVIRGRILNNGAYSLTLOGISWLGRLSRLSGSLALIHNTHLCFVHTV 480
DB 421 DLSVFQRLQVIRGRILNNGAYSLTLOGISWLGRLSRLSGSLALIHNTHLCFVHTV 480
QY 481 PMDQLPNNPQALHTANRPEDCEVSGELACHOLCARGHGCMGPPTQCVNCSOPLRGQEC 540
DB 481 PMDQLPNNPQALHTANRPEDCEVSGELACHOLCARGHGCMGPPTQCVNCSOPLRGQEC 540
QY 541 VESCRVLQGI.PREYVNAARHCLPCHPECOPONGSTTCGPRADQCVACATKDPFCYARC 600
DB 541 VESCRVLQGI.PREYVNAARHCLPCHPECOPONGSTTCGPRADQCVACATKDPFCYARC 600
QY 601 PSGVKPDLSTMPIMKPPDEBGACOPCPINCHSCVDLDDKCPAPORASPLTISIISAVG 660
DB 601 PSGVKPDLSTMPIMKPPDEBGACOPCPINCHSCVDLDDKCPAPORASPLTISIISAVG 660
QY 661 ILAVVVGAVVGLIKRROQKIRKTYMRRLQETELVEPLTPSGAMPNOQMRILKTEHL 720
DB 661 ILAVVVGAVVGLIKRROQKIRKTYMRRLQETELVEPLTPSGAMPNOQMRILKTEHL 720
QY 721 RKVAVLGSAGFTYKGIWIPDGENVKI.PVAIKYLRENTSFKANKEILDEAYVMAVGSP 780
DB 721 RKVAVLGSAGFTYKGIWIPDGENVKI.PVAIKYLRENTSFKANKEILDEAYVMAVGSP 780
QY 781 YVSLHLDGICLSTVQVLTQMLPVGCLLDHYENNGRLGSDOLLMWCMQIAGKMSYLDVR 840
DB 781 YVSLHLDGICLSTVQVLTQMLPVGCLLDHYENNGRLGSDOLLMWCMQIAGKMSYLDVR 840

QY 841 LVHRDIAANVLYKSNHYKTFDGLARLLIDETSYHADGKVPKIMMALESTLRRRT 900
DB 841 LVHRDIAANVLYKSNHYKTFDGLARLLIDETSYHADGKVPKIMMALESTLRRRT 900
QY 901 HQSDVNSYGVYWEIMTFPAKPYDGI.PABEIPDLKEGRLPDPPICTIVYIMVNCMM 960
DB 901 HQSDVNSYGVYWEIMTFPAKPYDGI.PABEIPDLKEGRLPDPPICTIVYIMVNCMM 960
QY 961 ISECRPRELVESESRMARDPQRFVIONEDLGPASPLDSTFYRSILBDDMGDLVDA 1020
DB 961 ISECRPRELVESESRMARDPQRFVIONEDLGPASPLDSTFYRSILBDDMGDLVDA 1020
QY 1021 ERYLVPOQGFPCDPAPAGGVVHHRSSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
DB 1021 ERYLVPOQGFPCDPAPAGGVVHHRSSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
QY 1081 AGSDVFDGDLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVAPLTCSPQREYV 1140
DB 1081 AGSDVFDGDLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVAPLTCSPQREYV 1140
QY 1141 NQPDVAPQPPSPREGPLPAPAPAGATLERPKTISPKNQGVKDVAFGAVENPEYLTPQ 1200
DB 1141 NQPDVAPQPPSPREGPLPAPAPAGATLERPKTISPKNQGVKDVAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPPERGAPSPSTFKGTPTABNPEYLGIDPV 1255
DB 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPPERGAPSPSTFKGTPTABNPEYLGIDPV 1255

RESULT 4
AAV84780
ID AAV84780 standard; Protein; 1255 AA.
XX
AC AAV84780;
XX
DT 08-ANG-2000 (first entry)
XX
DE Amino acid sequence of the SPLICE erbB-2 receptor protein.

XX
XX SPLICE erbB-2 receptor protein; cell transformation disorder; cancer;
KW tumor cell proliferation; tissue degeneration; atrophy;
KW bone resorption; inflammatory disease; degenerative disorder;
KW wound healing.
OS Homo sapiens.
XX
PN NC0200020579-A1.
PD 13-APR-2000.
XX
XX 01-OCT-1999; 99MO-CA00912.
XX PF
XX 02-OCT-1998; 98US-0165192.
XX
PA (UIMC-) UNIV MCMASTER.
XX
PI Muller WJ, Siegel PM;
XX
XX WPI, 2000-103768/26.
DR N-PSDB; AAA14812.
XX
XX

PT Nucleic acid encoding an erbB-2 receptor protein designated SPLICE
PT erbB-2, inhibitors of the protein are useful for treatment of cancer -
XX
XX
PS Claim 3; Fig 2; 60pp; English.

CC The present sequence represents a SPLICE erbB-2 receptor protein. The
CC protein has an in-frame deletion of 16 amino acids, 2 of which are
CC conserved cysteine residues, compared to the unspliced protein. The
CC erbB-2 polynucleotide is used to construct probes for detecting
CC disorders of cell transformation such as cancer. Antibodies to the
CC protein may be used to detect SPLICE erbB-2 in a sample. Agents
CC (e.g. antisense oligonucleotides) which inhibit the expression of

CC Splice erbb-2 are useful for reducing tumor cell proliferation and
 CC treating cancer. Substances which stimulate Splice erbb-2 are useful
 CC for treating conditions involving damaged cells including conditions
 CC in which degeneration of tissue occurs, such as arthropathy, bone
 CC resorption, inflammatory diseases, degenerative disorders of the
 CC central nervous system and wound healing.

XX Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 21; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALCRWGLALLPPGAASSTOYCTGTMTKALPASPEETHLDMRLHLYGCGVQGNL 60
 DB 1 MELALCRWGLALLPPGAASSTOYCTGTMTKALPASPEETHLDMRLHLYGCGVQGNL 60
 QY 61 ELTYLPTNASLSPLODIOEVGVYLIAHNOVROVPLQRLIRVGTOLPEDVYALAVDNG 120
 DB 61 ELTYLPTNASLSPLODIOEVGVYLIAHNOVROVPLQRLIRVGTOLPEDVYALAVDNG 120
 QY 121 DPLNNTPTVGASPGRLRLQRLSTELIKGVLIQRNPOLCYODTILMKDIPIKNNOLA 180
 DB 121 DPLNNTPTVGASPGRLRLQRLSTELIKGVLIQRNPOLCYODTILMKDIPIKNNOLA 180
 QY 181 LTLIDTNRSPACHSCPMCKSGRCMGSSESDQSLTRTVAGGACRCKPLPTDCHEQC 240
 DB 181 LTLIDTNRSPACHSCPMCKSGRCMGSSESDQSLTRTVAGGACRCKPLPTDCHEQC 240
 QY 241 AAGCTGKHSDDCLALPHNSGICELHCPALVTNTDTFESMNPBEGRYTFGASCTTAC 300
 DB 241 AAGCTGKHSDDCLALPHNSGICELHCPALVTNTDTFESMNPBEGRYTFGASCTTAC 300
 QY 301 YNTLSTVGSCTIYVCPHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
 DB 301 YNTLSTVGSCTIYVCPHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVAATVSAN 360
 QY 361 IOEPAGCKKI FGSIALFLPESFDGPASNTAPLOEOLQVETLEITGYLYISAMPDLP 420
 DB 361 IOEPAGCKKI FGSIALFLPESFDGPASNTAPLOEOLQVETLEITGYLYISAMPDLP 420
 QY 421 DLSVPQOLQYIRRIILNGAYSLTLOGLISMLGLRLRELSGLALIHNTHLCPVHTV 480
 DB 421 DLSVPQOLQYIRRIILNGAYSLTLOGLISMLGLRLRELSGLALIHNTHLCPVHTV 480
 QY 481 PMDOLFNNPHQALLHTANRDEDCVSGELACHOLCARGCMGPGPTQCVCNCSOFLRQEC 540
 DB 481 PMDOLFNNPHQALLHTANRDEDCVSGELACHOLCARGCMGPGPTQCVCNCSOFLRQEC 540
 QY 541 VEECRVLOGLPREYVNAHRLCPCHPCQOPNGSVTCFGEPAQOCVACAHYKDPFCVARC 600
 DB 541 VEECRVLOGLPREYVNAHRLCPCHPCQOPNGSVTCFGEPAQOCVACAHYKDPFCVARC 600
 QY 601 PEGVAKDLSMPYWKPFDEGACQPCPINCTHSCVDLIDKGCAPAEQASPLTISIAYVG 660
 DB 601 PEGVAKDLSMPYWKPFDEGACQPCPINCTHSCVDLIDKGCAPAEQASPLTISIAYVG 660
 QY 661 ILLVVLGVVFGILIKRROOKIRYTMRLLOETELVEPLTPSGAMNOAMRLKETEL 720
 DB 661 ILLVVLGVVFGILIKRROOKIRYTMRLLOETELVEPLTPSGAMNOAMRLKETEL 720
 QY 721 RKVKVLSGAGFYVYKGIWIPDGENVKIPVALIKVRENTSPKANKELIDRAYVWAGGSP 780
 DB 721 RKVKVLSGAGFYVYKGIWIPDGENVKIPVALIKVRENTSPKANKELIDRAYVWAGGSP 780
 QY 781 YVSRLLGICLTSTVOVLTQMLPGCLLDHYRENRKRLGSDLLNMCQIAGKMSYEDVR 840
 DB 781 YVSRLLGICLTSTVOVLTQMLPGCLLDHYRENRKRLGSDLLNMCQIAGKMSYEDVR 840
 QY 841 LVHRDLAARNVLVKSPHVKITDPRGLARLLDIDETEHADGKVPIMMALLESILRRFT 900
 DB 841 LVHRDLAARNVLVKSPHVKITDPRGLARLLDIDETEHADGKVPIMMALLESILRRFT 900

QY 901 HQSDVMSGYVTWELMTFGAKPYDGIPIAREIPDLLKGERLPOPEICTIDVYMINVKCM 960
 DB 901 HQSDVMSGYVTWELMTFGAKPYDGIPIAREIPDLLKGERLPOPEICTIDVYMINVKCM 960
 QY 961 IDSECRPRFRLVSEFSRMARDPQRFVYIQWEDJGSPASPLDSTFYSRLIEDDMGDLYDA 1020
 DB 961 IDSECRPRFRLVSEFSRMARDPQRFVYIQWEDJGSPASPLDSTFYSRLIEDDMGDLYDA 1020
 QY 1021 EBYLVPOQGFPCPPAGAGGVHRRSSSTRSGGGDLTGLFSESEAPSPAPSEB 1080
 DB 1021 EBYLVPOQGFPCPPAGAGGVHRRSSSTRSGGGDLTGLFSESEAPSPAPSEB 1080
 QY 1081 AGSDVPDGLGMAKAGLQSLPTHDPSPLOKYSDDPTVLPSETDGVYAPLTCSQPEV 1140
 DB 1081 AGSDVPDGLGMAKAGLQSLPTHDPSPLOKYSDDPTVLPSETDGVYAPLTCSQPEV 1140
 QY 1141 NQPDVRQPPSPRSGPLPAARPAATLERPKTSLPGKNGVYKDYAFGAVENPEYLTPO 1200
 DB 1141 NQPDVRQPPSPRSGPLPAARPAATLERPKTSLPGKNGVYKDYAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPPPAPSPAFDNLYYMDQDPBRGAPSTFGPTAENPEYLGIDVPV 1255
 DB 1201 GGAAPQHPPPAPSPAFDNLYYMDQDPBRGAPSTFGPTAENPEYLGIDVPV 1255

RESULT 5
 ID AAB85458 standard; protein; 1255 AA.
 XX AAB85458;
 AC AAB85458;
 DT 25-SEP-2001 (first entry)
 XX
 DE Human HER-2/neu protein.
 XX
 KW Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
 KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2.
 XX
 OS Homo sapiens.
 XX
 PN W0200153463-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001; 2001MO-US01850.
 XX
 PR 21-JAN-2000; 2000US-0177545.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Cheever MA, Hand-Zimmermann S;
 XX
 DR WPI, 2001-476112/51.
 DR N-PSDB; AAH23392.
 XX
 PT New antigen-presenting cells, useful as vaccines for eliciting or
 PT enhancing an immune response to HER-2/neu protein, particularly useful
 PT for treating or preventing cancer, e.g. breast cancer
 XX
 PS Claim 2; Page 41-46; 49pp; English.
 XX
 CC The invention provides an isolated antigen-presenting cell, which
 CC expresses at least an immunogenic portion of a polypeptide that produces
 CC an immune response to HER-2/neu protein. The antigen-presenting cells are
 CC useful as vaccines for eliciting or enhancing an immune response to
 CC HER-2/neu protein, particularly in treating or preventing malignancies in
 CC which the HER-2/neu oncogene is associated. Specifically, these are
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
 CC colon, lung or prostate cancers. The present sequence represents
 CC the human HER-2/neu protein (also known as p185 or c-erbB2).
 XX
 SQ Sequence 1255 AA;

CC vaccines provides the opportunity to combine epitopes derived from
 CC multiple tumour-associated molecules addressing the problem of tumour-
 CC tumour variability and reducing the likelihood of tumour escape due to
 CC antigen loss. AAG88266 to AAG89121 represent amino acid sequences used in
 CC the exemplification of the present invention.

Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 22; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 METALACRMGLLALIPGAASTQVCTGDMKRLPASPEHMLDMHLVQGCQVQGNL 60
 1 MELALACRMGLLALIPGAASTQVCTGDMKRLPASPEHMLDMHLVQGCQVQGNL 60
 61 ELTYLPTNLSLFLQDIQEVGYVLIANQVRQVPLRLIRVGTQLFEDNYALAVDNG 120
 61 ELTYLPTNLSLFLQDIQEVGYVLIANQVRQVPLRLIRVGTQLFEDNYALAVDNG 120
 121 DPANNTTPTVAGAPGGLRLQLRLTILKGVLIQGNPOLCYQDTILMKDIFHKNQLA 180
 121 DPANNTTPTVAGAPGGLRLQLRLTILKGVLIQGNPOLCYQDTILMKDIFHKNQLA 180
 181 LTLIDNRSBACHPCSPMCKSGRCKWSSSDQSLRTTVCAAGCARCKGLPTDCCHEQC 240
 181 LTLIDNRSBACHPCSPMCKSGRCKWSSSDQSLRTTVCAAGCARCKGLPTDCCHEQC 240
 241 AAGCTGPKISDCLACLFHNSGICELHCPALVTYNTDTPSMNPBEGRYFGASCYACP 300
 241 AAGCTGPKISDCLACLFHNSGICELHCPALVTYNTDTPSMNPBEGRYFGASCYACP 300
 301 YNVLSTDVSGCTLVCPILNQEVTAEDGTORCEKCSKRCARVCTGLGMEHLREYAVTSAN 360
 301 YNVLSTDVSGCTLVCPILNQEVTAEDGTORCEKCSKRCARVCTGLGMEHLREYAVTSAN 360
 361 IQEPACCKTIFGLALPLPSFPDDPASNTPPLQEPQLQVETLEITGYLYISAMPDSL 420
 361 IQEPACCKTIFGLALPLPSFPDDPASNTPPLQEPQLQVETLEITGYLYISAMPDSL 420
 421 DLSVFQNLQVIRGRILANGAYSLTLQGLISWLGSLSLRLSGSLALIHNTLCEVHTV 480
 421 DLSVFQNLQVIRGRILANGAYSLTLQGLISWLGSLSLRLSGSLALIHNTLCEVHTV 480
 481 PWDOLFERNPHQALLHTANRPEDECVEGEGLAHQLCARGHCWGPPTQCVNCSQFLRGQEC 540
 481 PWDOLFERNPHQALLHTANRPEDECVEGEGLAHQLCARGHCWGPPTQCVNCSQFLRGQEC 540
 541 VEECRVLOGLPRBYVNAHCLPCHPECCQPNQSVTCFGEPAQCVA CAHYKDPFCVARC 600
 541 VEECRVLOGLPRBYVNAHCLPCHPECCQPNQSVTCFGEPAQCVA CAHYKDPFCVARC 600
 601 PSGVPRDLSYMPILWKPEDEEGACQPCINCTHSCVLDLDKGCAEGRASPLTISAVVG 660
 601 PSGVPRDLSYMPILWKPEDEEGACQPCINCTHSCVLDLDKGCAEGRASPLTISAVVG 660
 661 ILVVVVLGVVFGILLIRROOKIRKTYMRRLQETELVEPLTPSGAMPNOAKMLKETE 720
 661 ILVVVVLGVVFGILLIRROOKIRKTYMRRLQETELVEPLTPSGAMPNOAKMLKETE 720
 721 RKVKVLGSGAGFTVYKGIWIPDGENVKIFVAIKVLEBNTSPKANKELIDBAVYMAVGSP 780
 721 RKVKVLGSGAGFTVYKGIWIPDGENVKIFVAIKVLEBNTSPKANKELIDBAVYMAVGSP 780
 781 YVSRLLGICLTSTVQLVQLMPGCLLDVYREKRGGLSODLIANCMQIAKMSYLEDVR 840
 781 YVSRLLGICLTSTVQLVQLMPGCLLDVYREKRGGLSODLIANCMQIAKMSYLEDVR 840
 841 LVHRDLAARNVAVKSPNHYKITDFGLARLLDIDETETHADGGKVPKMMALSTILRRPT 900
 841 LVHRDLAARNVAVKSPNHYKITDFGLARLLDIDETETHADGGKVPKMMALSTILRRPT 900
 901 HOSDVMSYGVTVWELMTFGAKPYDGI PARBIDPLEKGERLPQPICTIDVYMIWKCMM 960

DB 901 HOSDVMSYGVTVWELMTFGAKPYDGI PARBIDPLEKGERLPQPICTIDVYMIWKCMM 960
 QY 961 IDSECRFRELVESEFRMARDPQRFVIONEDGPASPLDSTYRSLLEDDMDGLVDA 1020
 DB 961 IDSECRFRELVESEFRMARDPQRFVIONEDGPASPLDSTYRSLLEDDMDGLVDA 1020
 QY 1021 ERYLVPOQGFPCDPAPAGAGVHHRSSSTSGGADLTGLPSESEARSPPLAPSEG 1080
 DB 1021 ERYLVPOQGFPCDPAPAGAGVHHRSSSTSGGADLTGLPSESEARSPPLAPSEG 1080
 QY 1081 AGSDVDFDGLGKAAKGLQSLPTHDPSPQLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140
 DB 1081 AGSDVDFDGLGKAAKGLQSLPTHDPSPQLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140
 QY 1141 NQPDVPRPQPSRPGPLPARPAGATLERPKTISPGKNGVVKVFAFGAVENPEYLTPO 1200
 DB 1141 NQPDVPRPQPSRPGPLPARPAGATLERPKTISPGKNGVVKVFAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPPEPAPSPAFDNLVYWDOPPERGAPSTFKGTPTAENPEYLGADV 1255
 DB 1201 GGAAPQHPPEPAPSPAFDNLVYWDOPPERGAPSTFKGTPTAENPEYLGADV 1255

RESULT 7

AAE24067
 AAE24067 standard; Protein; 1255 AA.

AC AAE24067;
 DT 23-SEP-2002 (first entry)
 XX Human Her-2 protein.
 DE Human Her-2 protein.
 XX Human; Her-2; epidermal growth factor receptor 2; infection; cancer;
 KW hyperproliferative disorder; prophylaxis; inflammation; antisense;
 KM tumour; gene therapy; phosphorothioate backbone.
 XX Homo sapiens.

PN WO200222636-A1.
 XX 21-MAR-2002.
 PD 12-SEP-2001; 2001WO-US28572.
 PF 15-SEP-2000; 2000US-0663834.
 PR (ISIS-) ISIS PHARM INC.
 XX Bennett CF, Cowseart LM;
 PA WPI; 2002-471192/50.
 XX N-PSDB; AAD88904.
 DR N-PSDB; AAD88904.

PT Novel antisense oligonucleotide which modulates the expression of Human
 PT Epidermal Growth Factor receptor, Her2, is useful for treating tumors
 PT inflammation or to prevent infection in humans -
 XX Example 13; Page 95-107; 116pp; English.

CC The invention relates to antisense compounds targeted to a nucleic
 CC acid molecule encoding Her2 (human Epidermal Growth Factor receptor 2)
 CC that specifically hybridises with and inhibits the expression of Her2.
 CC Antisense compounds of the invention are used for treating diseases or
 CC conditions associated with Her2 such as hyperproliferative disorders
 CC e.g. lung, breast, gastric, oesophageal, colon, bladder, salivary,
 CC neural or cardiac cancer. They are also useful prophylactically e.g.
 CC to prevent or delay infection, inflammation and tumour formation. The
 CC invention is also used in gene therapy. The present sequence is human
 CC Her-2 protein.

Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 23; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MELALCRWGLLALPPGASVQVCTGTMTKRLPASPTHMLMLHLYOGCOVGNL 60
 1 MELALCRWGLLALPPGASVQVCTGTMTKRLPASPTHMLMLHLYOGCOVGNL 60
 61 ELTYLPTNASLSFLQDIQEVGYVLIANQVRQVPLRLRVSTQLEFBNYALAVDNG 120
 61 ELTYLPTNASLSFLQDIQEVGYVLIANQVRQVPLRLRVSTQLEFBNYALAVDNG 120
 121 DPLANTTPTVGSAGGRLRELOLSSTILKGVLIQNNPOLCYODTILKMDIFKNNOLA 180
 121 DPLANTTPTVGSAGGRLRELOLSSTILKGVLIQNNPOLCYODTILKMDIFKNNOLA 180
 181 LTLIDNRSRACHPCSPCKGSRGWSSESDCOSLTRTVACAGGACRCKPLPDCCHQEC 240
 181 LTLIDNRSRACHPCSPCKGSRGWSSESDCOSLTRTVACAGGACRCKPLPDCCHQEC 240
 241 AACCTGKSHDCLACLFHNSGICELHCPALVTYNTDTPESMPNBSRGYTFGASCTVACP 300
 241 AACCTGKSHDCLACLFHNSGICELHCPALVTYNTDTPESMPNBSRGYTFGASCTVACP 300
 301 YNLTSTDVSGCTVCPILHNOEVTAEQTCRCKSKPCANVCYGLGMEHLREVAVTSAN 360
 301 YNLTSTDVSGCTVCPILHNOEVTAEQTCRCKSKPCANVCYGLGMEHLREVAVTSAN 360
 361 IOEPAGCKKI FGSIAFLPESPPDGPASNTAPLOEBOLOVEETLEITGYLISMPDLP 420
 361 IOEPAGCKKI FGSIAFLPESPPDGPASNTAPLOEBOLOVEETLEITGYLISMPDLP 420
 421 DLSVFNQLQVIRGRILHNGAYSLTLOGLISWLGRLSRLSGLALIHNTHLCPHTV 480
 421 DLSVFNQLQVIRGRILHNGAYSLTLOGLISWLGRLSRLSGLALIHNTHLCPHTV 480
 481 PMDLFNPQHIALHTNRPBDECVGEGALACHOLCARGCKGPGPTCCVCSQPLRGQEC 540
 481 PMDLFNPQHIALHTNRPBDECVGEGALACHOLCARGCKGPGPTCCVCSQPLRGQEC 540
 541 VESCRVLOGLPREVYNARHCLPCHBEGOPONGSVTCGPRADQCVACAHKDPFCVARG 600
 541 VESCRVLOGLPREVYNARHCLPCHBEGOPONGSVTCGPRADQCVACAHKDPFCVARG 600
 601 PSQVPELSTYMPILKFPDEBEGACQPCPINCSTHCVDDLDDGCPAPQASPLTSLISAVG 660
 601 PSQVPELSTYMPILKFPDEBEGACQPCPINCSTHCVDDLDDGCPAPQASPLTSLISAVG 660
 661 ILLVVLGVVFGILLIKRQOKIRKRYNRLLQSTELVEPLTPSGAMPNOAMRLKETEL 720
 661 ILLVVLGVVFGILLIKRQOKIRKRYNRLLQSTELVEPLTPSGAMPNOAMRLKETEL 720
 721 RKRYVGGSGAGTYVYKGIWIPDGENVKIPVALIKYLRNTSPKAKELIDELVYVAGGSP 780
 721 RKRYVGGSGAGTYVYKGIWIPDGENVKIPVALIKYLRNTSPKAKELIDELVYVAGGSP 780
 781 YVSRLLGICLTSVQVLTQMLPFGCLDHYENRGRIGSODLNMCMQIAKMSYLEDVR 840
 781 YVSRLLGICLTSVQVLTQMLPFGCLDHYENRGRIGSODLNMCMQIAKMSYLEDVR 840
 841 LVHRDLAARVNLVKSPPHVKITDFGLARLLDIDETBYHADGKVPKIMMALBSILRRFT 900
 841 LVHRDLAARVNLVKSPPHVKITDFGLARLLDIDETBYHADGKVPKIMMALBSILRRFT 900
 901 HOSDVSYGTVVWELMTFGAKPYDGIARERLIDLEBGERLPPOPICITIDVYIMVACMM 960
 901 HOSDVSYGTVVWELMTFGAKPYDGIARERLIDLEBGERLPPOPICITIDVYIMVACMM 960
 961 IDSECRFRFELVSEFARMARDPQPVYIQNEDLGPASPLDSTYRSLLDEDDMGDLVDA 1020
 961 IDSECRFRFELVSEFARMARDPQPVYIQNEDLGPASPLDSTYRSLLDEDDMGDLVDA 1020

QY 1021 EELTVPOQGFPCDPAPAGAGVHHRSSSTRSGGDLTLGLRPSREAPRSLAPSEG 1080
 DB 1021 EELTVPOQGFPCDPAPAGAGVHHRSSSTRSGGDLTLGLRPSREAPRSLAPSEG 1080
 QY 1081 AGSDVFDGDLGMAAGLQSLPTHDPSPLOGRYSDDPTVPLPSETDGVVAPLTCSPOPEYV 1140
 DB 1081 AGSDVFDGDLGMAAGLQSLPTHDPSPLOGRYSDDPTVPLPSETDGVVAPLTCSPOPEYV 1140
 QY 1141 NQPDVAPQPPSPREGPLPAARPGATLERPKTSLPGKNGVVKOVFAFGAVENBEYLTPQ 1200
 DB 1141 NQPDVAPQPPSPREGPLPAARPGATLERPKTSLPGKNGVVKOVFAFGAVENBEYLTPQ 1200
 QY 1201 GGAAPQHPHPAPASPAFDNLTYWDOPPERGAPSTFKGPTAENPEYLGIDVAV 1255
 DB 1201 GGAAPQHPHPAPASPAFDNLTYWDOPPERGAPSTFKGPTAENPEYLGIDVAV 1255

RESULT 8
 ID AAE20479 standard; Protein; 1255 AA.
 AC AAE20479;
 DT 01-JUL-2002 (first entry)
 DE Human Her-2/neu protein.
 KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 1021..1030
 XX /note="Naturally processed HLA-B44-restricted epitope"
 XX NO200214503-A2.
 PD 21-FEB-2002.
 PF 14-AUG-2001; 2001MO-US41733.
 PR 14-AUG-2000; 2000US-225152P.
 PR 28-SEP-2000; 2000US-236428P.
 PR 21-FEB-2001; 2001US-270520P.
 PA (CORI-) CORIAX CORP.
 PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
 PI McNeill PD, Vedvick TS;
 DR WPI; 2002-280758/32.
 DR N-PSDB; AAD32743.
 PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 PT prevention and diagnosis of cancer, preferably breast cancer -
 XX
 XX Disclosure; Page 114-117; 129pp; English.
 CC The invention relates to an isolated Her-2/Neu polypeptide composition
 CC effective for eliciting an immune response. The invention is useful for
 CC eliciting an immune response in a patient, where the patient is human
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
 CC The composition is useful for the therapy and diagnosis of cancer,
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 CC and other compositions for the diagnosis, prevention and treatment of
 CC human malignancies, for stimulating and/or expanding T cells specific for
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 CC patient. The invention is useful for stimulating a T cell response in a
 CC human patient, as probe or primer for nucleic acid hybridisation, to
 CC selectively form duplex molecules with complementary stretches of the
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in
 CC prophylactic or therapeutic applications and for the treatment of cancer,
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
 CC associated malignancies. The invention is useful in gene therapy. The
 CC present sequence is human Her-2/neu protein.

XX Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 23; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METAAACRGGMLLALPPGAASVQVCTGTMTKRLPASPETHLDMLRHLYQGGVVOGNL 60
 DB 1 METAAACRGGMLLALPPGAASVQVCTGTMTKRLPASPETHLDMLRHLYQGGVVOGNL 60
 QY 61 ELTYLPTNASLSLFDIPIQEVGYVLIAMNOVROVPLORLRIIVRGTOLEFEDNYALAVDNG 120
 DB 61 ELTYLPTNASLSLFDIPIQEVGYVLIAMNOVROVPLORLRIIVRGTOLEFEDNYALAVDNG 120
 QY 121 DPLANTTPVTGASPGGLRELQLASLTELKGVLIQBNPOLCYQDTILMWDIFHKNNQLA 180
 DB 121 DPLANTTPVTGASPGGLRELQLASLTELKGVLIQBNPOLCYQDTILMWDIFHKNNQLA 180
 QY 181 LTLIDTRRSBAACPGSPCKSGRCCKGSSSEDCCSLRTTVCAAGCARCKGPLPTDCHEOC 240
 DB 181 LTLIDTRRSBAACPGSPCKSGRCCKGSSSEDCCSLRTTVCAAGCARCKGPLPTDCHEOC 240
 QY 241 AAGCTGPKKSDCLACHPNHSGLCELCAPALVYNTDTPESMPNBERATYFGASCVTACP 300
 DB 241 AAGCTGPKKSDCLACHPNHSGLCELCAPALVYNTDTPESMPNBERATYFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVCPDLNQEVTABDGTORCEKSKPCAVCYGLGMEHLREYRAVTSAN 360
 DB 301 YNYLSTDVGSCTLVCPDLNQEVTABDGTORCEKSKPCAVCYGLGMEHLREYRAVTSAN 360
 QY 361 IQGFACCKKIFGSLATLPESFDGDPASNTAPLOPEOLQVETTLBEITGYLYISAMPDSL 420
 DB 361 IQGFACCKKIFGSLATLPESFDGDPASNTAPLOPEOLQVETTLBEITGYLYISAMPDSL 420
 QY 421 DLSVFNQLVIRRIILHNGAYSLTLQGLISWLGSLRELSSGLALIHNTHLCPVHTV 480
 DB 421 DLSVFNQLVIRRIILHNGAYSLTLQGLISWLGSLRELSSGLALIHNTHLCPVHTV 480
 QY 481 PMDQLFRNPHQALHTANRPEDECVGBGLACHOLCARHGMCGAPTCVNCQPLRGQEC 540
 DB 481 PMDQLFRNPHQALHTANRPEDECVGBGLACHOLCARHGMCGAPTCVNCQPLRGQEC 540
 QY 541 VEECRVLQGLPREYVNAHRLCPHPCQPONGSVTCGPBADOCAAHYKDBPFCVARC 600
 DB 541 VEECRVLQGLPREYVNAHRLCPHPCQPONGSVTCGPBADOCAAHYKDBPFCVARC 600
 QY 601 BEGVKEDLSYMPYWKPPDEEGACQPCPINCTHSCVDLDKGCAPABASPLTIIISAVNG 660
 DB 601 BEGVKEDLSYMPYWKPPDEEGACQPCPINCTHSCVDLDKGCAPABASPLTIIISAVNG 660
 QY 661 ILLVVVLGVVFGILLIRROOKIRKYMBSRLQETELVEPLTPSGAMPNOMKILKTEL 720
 DB 661 ILLVVVLGVVFGILLIRROOKIRKYMBSRLQETELVEPLTPSGAMPNOMKILKTEL 720
 QY 721 RKVKVLSGSGFGTVYKGIWIPDEBNYKIPVALIYVLENTSPKANKELIDEAAYVAGVSP 780
 DB 721 RKVKVLSGSGFGTVYKGIWIPDEBNYKIPVALIYVLENTSPKANKELIDEAAYVAGVSP 780
 QY 781 YVSRLLGICLTSTVOLVTOQLMPYGLLDHRENRGRGLSGDLDLNMWCMQIAKMSYLEDVR 840
 DB 781 YVSRLLGICLTSTVOLVTOQLMPYGLLDHRENRGRGLSGDLDLNMWCMQIAKMSYLEDVR 840
 QY 841 LVHRDLAARNVLYKSPHNVKITDFGLARLLDIDETEHADGKVPYIKMMLJESILRRRPT 900
 DB 841 LVHRDLAARNVLYKSPHNVKITDFGLARLLDIDETEHADGKVPYIKMMLJESILRRRPT 900
 QY 901 HOSDVMISYGVVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPPICTIDVYIMVYKCM 960

DB 901 HOSDVMISYGVVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPPICTIDVYIMVYKCM 960
 QY 961 IDSECRPRPRELYSESRMARDDQRFVVIQNEDLGPASPLDSTFYRSLDEDDMDGLVDA 1020
 DB 961 IDSECRPRPRELYSESRMARDDQRFVVIQNEDLGPASPLDSTFYRSLDEDDMDGLVDA 1020
 QY 1021 EEYLVPOQGFCDPAPAGAGVHHRRSSSTRSGGDLTLGLEPSEEBAPRSPLASSEG 1080
 DB 1021 EEYLVPOQGFCDPAPAGAGVHHRRSSSTRSGGDLTLGLEPSEEBAPRSPLASSEG 1080
 QY 1081 AGSDVDGDLGMGAAGLQSLPTHDPSPLORYSEDPVLPSESTDGYVAPLTCSPOBEYV 1140
 DB 1081 AGSDVDGDLGMGAAGLQSLPTHDPSPLORYSEDPVLPSESTDGYVAPLTCSPOBEYV 1140
 QY 1141 NQPDVAPQPPSPREGPLPAARPGATLERPKTSPGKNGVXQVFAFGAVENPEYLTPQ 1200
 DB 1141 NQPDVAPQPPSPREGPLPAARPGATLERPKTSPGKNGVXQVFAFGAVENPEYLTPQ 1200
 QY 1201 GGAAPQHPHPPAPSPADNLVYWDQPPERGAPSTFKGTPTAENPEYLGIDVAV 1255
 DB 1201 GGAAPQHPHPPAPSPADNLVYWDQPPERGAPSTFKGTPTAENPEYLGIDVAV 1255

RESULT 9
 AAM51143
 ID AAM51143 standard; Protein; 1255 AA.
 XX
 AC AAM51143;
 XX
 DT 17-JUN-2002 (first entry)
 XX
 DE Human Her-2/neu oncogene-encoded p185 glycoprotein.
 XX
 KM Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
 OS Tyrosine kinase; receptor; c-erbB2; gene therapy.
 OS Homo sapiens.
 XX
 FH Key
 FT Domain 1..653 Location/Qualifiers
 FT Domain /note="extracellular domain"
 FT Domain 676..1255
 FT Domain /note="intracellular domain"
 FT Domain 990..1255
 FT Domain /note="phosphorylation domain"
 XX
 EN WO200212341-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 03-AUG-2001; 2001MO-US24283.
 XX
 PR 03-AUG-2000; 2000US-0632507.
 XX
 PA (CORI-) CORIXA CORP.
 PA (SMIK-) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Cheever MA, Cheysen D;
 XX
 DR WPI; 2002-241743/29.
 DR N-PSDB; ABA02250.
 XX
 PT Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or
 PT phosphorylation domain
 XX
 PS Claim 68; Fig 7; 14pp; English.
 XX
 CC The present sequence is that of human Her-2/neu (p185 glycoprotein
 CC or c-erbB2), an oncogenic self-protein and target for anti-cancer
 CC vaccine. The Her-2/neu gene is amplified and p185 is overexpressed

in a variety of cancers, including breast, ovarian, colon, lung and prostate cancer. Her-2/neu is a member of the tyrosine kinase family of receptor-like glycoproteins. It comprises an extracellular domain with homology to the epidermal growth factor receptor (EGFR), a highly hydrophobic transmembrane domain and a C-terminal intracellular domain that also shows homology to EGFR. Its overexpression correlates with a poor prognosis in breast and ovarian cancers. The invention provides Her-2/neu fusion proteins, nucleic acids encoding them, viral vectors, and vaccines comprising the fusion proteins or nucleic acid molecules. In preferred fusion proteins, the extracellular domain of a Her-2/neu protein is fused to a Her-2/neu intracellular domain or phosphorylation domain (or its deleted fragment). An immune response to Her-2/neu protein is elicited or enhanced by administering the fusion protein in the form of a vaccine, or by transfecting cells of an animal *ex vivo* with a nucleic acid encoding the fusion protein, and delivering the transfected cells to the animal. The fusion proteins, nucleic acids, and isolated specific T-cells are useful for inhibiting the development of a cancer, especially breast, ovarian, colon, lung or prostate cancer in a patient. T cells that specifically react with a Her-2/neu fusion protein can be used to remove tumour cells from a sample in order to inhibit the development of cancer in a patient.

Sequence	1255 AA;
Query Match	100.0%; Score 6815; DB 23; Length 1255;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1255; Conservative	0; Mismatches 0; Indels 0; Gaps 0

Qy	1	MELALACRMGLLALLPCGASTVCOTGTDMLRLPASETHLDMRLHYOGCQVVGSL	60
Db	1	MELALACRMGLLALLPCGASTVCOTGTDMLRLPASETHLDMRLHYOGCQVVGSL	60
Qy	61	ELTYLPTNASLFLDDIQEVSQYLIAHQVQVPLQRLRYRGQLPREDNALVLNG	120
Db	61	ELTYLPTNASLFLDDIQEVSQYLIAHQVQVPLQRLRYRGQLPREDNALVLNG	120
Qy	121	DLPLNTTPTVGTASPGGLRELOLRSLTEILKGVLLIQRNPOLCQOTIIMKDIIFHKNQOLA	180
Db	121	DLPLNTTPTVGTASPGGLRELOLRSLTEILKGVLLIQRNPOLCQOTIIMKDIIFHKNQOLA	180
Qy	181	LTLITPNSRACHPESPMCKSGRCMGESSEDDQSILTRYCAGGACRCRKLPTDCCHEQC	240
Db	181	LTLITPNSRACHPESPMCKSGRCMGESSEDDQSILTRYCAGGACRCRKLPTDCCHEQC	240
Qy	241	AAGCTGPKHSDCLACLFPHSHGICELHCALVTYNTDTRESMPNEGRTTPASCVTACP	300
Db	241	AAGCTGPKHSDCLACLFPHSHGICELHCALVTYNTDTRESMPNEGRTTPASCVTACP	300
Qy	301	YNYLSTDVGSCTLYCPLHQVLTABDGTQRCEKCSKPCARVCYGLAMEHLREVRAVTSAN	360
Db	301	YNYLSTDVGSCTLYCPLHQVLTABDGTQRCEKCSKPCARVCYGLAMEHLREVRAVTSAN	360
Qy	361	IOEPAGCKKIPESLAFLEPSPGDGPRASNPALQPBOLQVFTLEBETLGYLIYANPDSLP	420
Db	361	IOEPAGCKKIPESLAFLEPSPGDGPRASNPALQPBOLQVFTLEBETLGYLIYANPDSLP	420
Qy	421	DLSPFONLQVIRGRILHNGVSLTLQSLGISWLGRLSLELSGLALIHNNHLCFVHTV	480
Db	421	DLSPFONLQVIRGRILHNGVSLTLQSLGISWLGRLSLELSGLALIHNNHLCFVHTV	480
Qy	481	PMDOLPRNPHOLLHTANRPEDECVGEBGLACHOLCARGCMKBPGRQCNCSOFLRGDSC	540
Db	481	PMDOLPRNPHOLLHTANRPEDECVGEBGLACHOLCARGCMKBPGRQCNCSOFLRGDSC	540
Qy	541	VEECVTLQGLPRBYVNARHCLPCHPECPONGSVTCTGFEZADQVACAHYKDPFPCVANC	600
Db	541	VEECVTLQGLPRBYVNARHCLPCHPECPONGSVTCTGFEZADQVACAHYKDPFPCVANC	600
Qy	601	PGGVKPRDLSYMPYWKPRDBEGACQCPINCTHSCVDLDDKGCPRABERASPLTISIAYVG	660
Db	601	PGGVKPRDLSYMPYWKPRDBEGACQCPINCTHSCVDLDDKGCPRABERASPLTISIAYVG	660

Qy	661	ILVWVUGVYFGLIKRROOKIKRMYRRLLOETHELVEPLTPSGAMNOMOMRLIKETEL	720
Db	661	ILLVWVGAVFGLIKRROOKIKRMYRRLLOETHELVEPLTPSGAMPNOAMRLIKETEL	720
Qy	721	RKVYVLSGSAFGVYKSIWI PDGENVYK PAI KYLRENTSPKANKELIDEYVWAGVSP	780
Db	721	RKVYVLSGSAFGVYKSIWI PDGENVYK PAI KYLRENTSPKANKELIDEYVWAGVSP	780
Qy	781	YVSLBLGICLTSTVOLVTO LMPYGCLLDHVENRGLSGODLLWMCQIAKMSYLEDVR	840
Db	781	YVSLBLGICLTSTVOLVTO LMPYGCLLDHVENRGLSGODLLWMCQIAKMSYLEDVR	840
Qy	841	LYVHEDLAARNVLYVSPHNKYITDGBGLARLIDIDETETHAOGGKPIKMMALESI LRERFT	900
Db	841	LYVHEDLAARNVLYVSPHNKYITDGBGLARLIDIDETETHAOGGKPIKMMALESI LRERFT	900
Qy	901	HOSDWSYGVYVWELMTFGAKPYDGI PARBI PDLLEKGERLPPOPPI CTIDVYMIWVCKM	960
Db	901	HOSDWSYGVYVWELMTFGAKPYDGI PARBI PDLLEKGERLPPOPPI CTIDVYMIWVCKM	960
Qy	961	IDSCRPFRFELVSEFSRMARDPQRFVYI QWEDLGSPAS PLDSTFYRSILLEDDMDGLVDA	1020
Db	961	IDSCRPFRFELVSEFSRMARDPQRFVYI QWEDLGSPAS PLDSTFYRSILLEDDMDGLVDA	1020
Qy	1021	EZYLYPQOGPFCCPDPA PGAGMWHRRRSSSTRSGGDDL TLGLPSEEBEARSP LABSEG	1080
Db	1021	EZYLYPQOGPFCCPDPA PGAGMWHRRRSSSTRSGGDDL TLGLPSEEBEARSP LABSEG	1080
Qy	1081	AGSDVFPQDGLGMAKAGLQSLPTHDPS PLQVYSSDPVYPLPSEYTDGVYAP LTCSPQPEYV	1140
Db	1081	AGSDVFPQDGLGMAKAGLQSLPTHDPS PLQVYSSDPVYPLPSEYTDGVYAP LTCSPQPEYV	1140
Qy	1141	NQPDVRPOPSPREGPI PAARPA GATLERPTTSLPGKNGVYKDVFAFGAVENDEYLTPO	1200
Db	1141	NQPDVRPOPSPREGPI PAARPA GATLERPTTSLPGKNGVYKDVFAFGAVENDEYLTPO	1200
Qy	1201	GGAAPQHPHPAPASPAEDNL YWQODDPERKAPSTFGCTANPEYLGIDVPU	1255
Db	1201	GGAAPQHPHPAPASPAEDNL YWQODDPERKAPSTFGCTANPEYLGIDVPU	1255

RESULT 10		
AAU77114		
ID	AAU77114	standard; Protein: 1255 AA.
XX		
AC	AAU77114;	
XX		
DT	05-JUN-2002	(first entry)
XX		
DE	Human Her-2/neu polypeptide.	
XX		
KW	Human; Her2/neu; cytosol; haematological malignancy; CML;	
KW	acute myelogenous leukaemia; ALL; chronic myelogenous leukaemia; CLL;	
KW	chronic lymphocytic leukaemia; myeloma; non-Hodgkin's lymphoma; MDS;	
KW	Hodgkin's lymphoma; T cell therapy.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200213847-A2.	
XX		
PD	21-FEB-2002.	
XX		
PF	13-AUG-2001; 2001MO-US25408.	
XX		
PR	14-AUG-2000; 2000US-0638280.	
XX		
PR	28-SEP-2000; 2000US-0675904.	
XX		
PA	(CORI-) CORIXA CORP.	
XX		
PI	Gaiger A, Cheever MA, Hand-zimmermann S;	
XX		
DR	WPI; 280741/32.	

DR N-PSDB, ABK10730.
 XX Inhibiting haematological malignancy development by administering
 PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide
 PT encoding the polypeptide, or antigen presenting cells expressing the
 PT polypeptide
 XX Disclosure, Page 71-74, 74pp; English.
 PS The invention relates to a method for inhibiting development of
 CC haematological malignancy in a patient by administering a polypeptide
 CC comprising an immunogenic portion of Her-2/neu or a polynucleotide
 CC encoding the polypeptide. Antigen presenting cells that express the
 CC protein can also be administered. The sequences are used for inhibiting
 CC development of haematological malignancy such as acute myelogenous
 CC leukemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic
 CC leukemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's
 CC lymphoma. This sequence represents the human Her-2/neu polypeptide.
 CC
 SQ Sequence 1255 AA;
 Query Match 100.0%; Score 6815; DB 23; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEIALACRMGLIALLPFGAASITQVCTGDMKRLPASPETHLMDLRLHYOGCQVQGNL 60
 DB 1 MEIALACRMGLIALLPFGAASITQVCTGDMKRLPASPETHLMDLRLHYOGCQVQGNL 60
 QY 61 ELTYLPTNLSLSTLOIQVQVYLAAHNOVROPRLRLIYNGTQLPENYALAVLDNG 120
 DB 61 ELTYLPTNLSLSTLOIQVQVYLAAHNOVROPRLRLIYNGTQLPENYALAVLDNG 120
 QY 121 DPLANTTPVTGASPGGLRELQSLRSLTEILKGVLIQBNPOLCYODITLMKDIHKNQOLA 180
 DB 121 DPLANTTPVTGASPGGLRELQSLRSLTEILKGVLIQBNPOLCYODITLMKDIHKNQOLA 180
 QY 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLRTTVACGAGCARCKGPLPTDCHEQC 240
 DB 181 LTLIDNRSRACHPCSPMKSGSRCKWSSSDCOSLRTTVACGAGCARCKGPLPTDCHEQC 240
 QY 241 AAGCTGPKASDCLACHFNHSGICELHCPALVYNTDTFESMNPBSGRYTFGASCTTAC 300
 DB 241 AAGCTGPKASDCLACHFNHSGICELHCPALVYNTDTFESMNPBSGRYTFGASCTTAC 300
 QY 301 YNYLSTDVSGCTVCPHNOEYTAADGTORCEKSKPCAVCYGLGMEHLREYRAVTSAN 360
 DB 301 YNYLSTDVSGCTVCPHNOEYTAADGTORCEKSKPCAVCYGLGMEHLREYRAVTSAN 360
 QY 361 IOEPACCKKI FGSIAFLPESFDDPASNTPALQPEOLQVETLEITGYLIISAMPDLP 420
 DB 361 IOEPACCKKI FGSIAFLPESFDDPASNTPALQPEOLQVETLEITGYLIISAMPDLP 420
 QY 421 DLSVFONLOVIRGRIILHNGAYSLTLOGLGISWLSLSLRELSGGLALIHNTHLCPVHTV 480
 DB 421 DLSVFONLOVIRGRIILHNGAYSLTLOGLGISWLSLSLRELSGGLALIHNTHLCPVHTV 480
 QY 481 PMDQLFRNPHOALLHTANREDECVBEGALCHOLCARGHWGSGPPQCVNCSQFLGQEC 540
 DB 481 PMDQLFRNPHOALLHTANREDECVBEGALCHOLCARGHWGSGPPQCVNCSQFLGQEC 540
 QY 541 VEECRVLOGLPREYVNAHCLPCGPECQPPNGSVTCGFPADQCVACAHYKDPFCVARC 600
 DB 541 VEECRVLOGLPREYVNAHCLPCGPECQPPNGSVTCGFPADQCVACAHYKDPFCVARC 600
 QY 601 PEGVKEPDLSTMPYWKTPDEEGACQPCPINCTHS CVLDLDDKCPAEPORASPLTIIISAVG 660
 DB 601 PEGVKEPDLSTMPYWKTPDEEGACQPCPINCTHS CVLDLDDKCPAEPORASPLTIIISAVG 660
 QY 661 ILLVVLGVVFGILLIRBROOKIRKTYMRLLQETELVBEPLTPSGAMPNQAOMRILKETEL 720
 DB 661 ILLVVLGVVFGILLIRBROOKIRKTYMRLLQETELVBEPLTPSGAMPNQAOMRILKETEL 720

QY 721 RKVYVLSGAGFVYKGIWIPDGENVKI PVAIKVLRNNTSPKANKEILDEAYVAGVSP 780
 DB 721 RKVYVLSGAGFVYKGIWIPDGENVKI PVAIKVLRNNTSPKANKEILDEAYVAGVSP 780
 QY 781 YSRLLIGICITSTVOLYTOAMPGLCLDHYRNRGRIGSDLLWMCQIAGMSYLEVR 840
 DB 781 YSRLLIGICITSTVOLYTOAMPGLCLDHYRNRGRIGSDLLWMCQIAGMSYLEVR 840
 QY 841 LVHRDLAARNVLYKSPNHVKITDGLARLLDIDETEHADGKVPKIMMALESILRRFT 900
 DB 841 LVHRDLAARNVLYKSPNHVKITDGLARLLDIDETEHADGKVPKIMMALESILRRFT 900
 QY 901 HQSDVMSYGVYVWELMTFGAKPYDGI PAREIPDLLEKGERLPDPICITIDVYMIWVKCM 960
 DB 901 HQSDVMSYGVYVWELMTFGAKPYDGI PAREIPDLLEKGERLPDPICITIDVYMIWVKCM 960
 QY 961 IDSECPREFELVSESRMARDPQRFVVIQNEBGLGPAAPLDSITFRSLLEDDMDGLVDA 1020
 DB 961 IDSECPREFELVSESRMARDPQRFVVIQNEBGLGPAAPLDSITFRSLLEDDMDGLVDA 1020
 QY 1021 EETLVPOQGFPCDPAPAGAMVHHRSSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
 DB 1021 EETLVPOQGFPCDPAPAGAMVHHRSSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
 QY 1081 AGSDVFDGDLGMAAGLOSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPOPEYV 1140
 DB 1081 AGSDVFDGDLGMAAGLOSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPOPEYV 1140
 QY 1141 NOPDVPPOPSPREGPLPAARPAATLERPKTISPGKNGYVKOVFAAGAVENBEYLTPQ 1200
 DB 1141 NOPDVPPOPSPREGPLPAARPAATLERPKTISPGKNGYVKOVFAAGAVENBEYLTPQ 1200
 QY 1201 GGAAPQHPPPAPAFDNLXYMDOPPERGAPSTKGTPTAENPEYLGLDVAV 1255
 DB 1201 GGAAPQHPPPAPAFDNLXYMDOPPERGAPSTKGTPTAENPEYLGLDVAV 1255

RESULT 11
 AA92620
 ID AA92620 standard; Protein; 1255 AA.
 XX
 AC AA92620;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human heregulin 2 (Her2).
 XX
 KW Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
 KM self-protein; cancer; breast cancer; prostate cancer;
 XX cell-associated peptide antigen; foreign epitope.
 OS Homo sapiens.
 XX
 FH Key
 FT Domain
 FT
 FT Location/Qualifiers
 FT 1..173
 FT /label= N-terminal
 FT /note= "mature polypeptide"
 FT 5..25
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 59..73
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 103..117
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 149..163
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 174..323
 FT /label= Cysteine-rich domain
 FT 210..224
 FT /label= insertion region

PT /note= "suitable for foreign epitope insertion"
 FT 250..264
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 324..483
 FT /label= ligand_binding_domain
 FT 325..339
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 369..383
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 465..479
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 484..623
 FT /label= Cysteine_rich_domain
 FT 579..593
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 624..654
 FT /label= Transmembrane_domain
 FT 632..652
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 653..667
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 655..1010
 FT /label= Tyrosine_kinase_domain
 FT 661..675
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 695..709
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 710..730
 FT /label= insertion region
 FT /note= "suitable for foreign epitope insertion"
 FT 1011..1235
 FT /label= C-terminal_domain
 FT Domain
 PN MO200020027-A2.
 XX 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WC-DK00525.
 XX
 PR 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX
 PA (MEBI-) M & B BIOTECH AS.
 PI Steinaa L, Mouritzen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX WPI, 2000-349917/30.
 DR N-PSDB; AAA09455.
 XX
 PT Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 PS
 XX Claim 62, Page 193-198, 220pp; English.
 CC This is the human heregulin 2 (Her2) sequence. Immunogenic analogues of
 CC Her2 can be used in the claimed method as an autovaccine to induce a CTL
 CC response. Subdominant CTL epitopes, antibody binding regions and
 CC cysteine residues involved in disulfide bonds are preserved in the
 CC immunogenized forms. Regions suitable for the insertion of foreign T
 CC helper epitopes were identified (see features table). The method
 CC is used for inducing immune responses against weakly immunogenic
 CC cell-associated peptide antigens (PA) such as those associated with
 CC cancers (self-proteins), e.g. human prostate specific membrane antigen

CC (PSM), heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b).
 CC The method comprises effecting simultaneous presentation by antigen
 CC producing cells (APCs) of the animal's immune system of: (1) at least 1
 CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the respective
 CC PA and including at least one foreign T helper epitope are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer
 CC when the PA is human PSM, FGF8b and Her2, respectively.
 XX
 SQ Sequence 1255 AA;
 Query Match 99.9%; Score 6806; DB 21; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MHAALCRWGLTALALPPGAASSTOVCTGDMKRLPASPTHTLMDLRHLYOGCGVQGNL 60
 DB 1 MHAALCRWGLTALALPPGAASSTOVCTGDMKRLPASPTHTLMDLRHLYOGCGVQGNL 60
 QY 61 ELTYLPTNASLSTLODIOEVQGYVLAHNOVRQVPLQRRLIVRGTOLEFBNYALAVLDNG 120
 DB 61 ELTYLPTNASLSTLODIOEVQGYVLAHNOVRQVPLQRRLIVRGTOLEFBNYALAVLDNG 120
 QY 121 DELNNTTPTVTSAPGGLRELQLRLSLRLKGVLIQRNPOLCOTDTLMDIFPKNNOLA 180
 DB 121 DELNNTTPTVTSAPGGLRELQLRLSLRLKGVLIQRNPOLCOTDTLMDIFPKNNOLA 180
 QY 181 LTLIDNRSRACHPCSPCKSGRCKWSSSEDCSLFTTVAGGACARCKPLPTCCHEQC 240
 DB 181 LTLIDNRSRACHPCSPCKSGRCKWSSSEDCSLFTTVAGGACARCKPLPTCCHEQC 240
 QY 241 AAGCTGPKSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTFGASCVTACP 300
 DB 241 AAGCTGPKSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTFGASCVTACP 300
 QY 301 YNLTSTDVGSCTLVCPHNOETABEDTORCEKSKRCANVCYGLGNEHLREVAVTSAN 360
 DB 301 YNLTSTDVGSCTLVCPHNOETABEDTORCEKSKRCANVCYGLGNEHLREVAVTSAN 360
 QY 361 IQEFAGCKKIFGSLAFPPESFPDGPASNTAPLOPEOLQVETLEIYGYIISWMPSLP 420
 DB 361 IQEFAGCKKIFGSLAFPPESFPDGPASNTAPLOPEOLQVETLEIYGYIISWMPSLP 420
 QY 421 DISVFQNLQVIRRIIHNAGVSLTLQGLGISMGLSLRLRIGSLALIHNTHLCPVHTV 480
 DB 421 DISVFQNLQVIRRIIHNAGVSLTLQGLGISMGLSLRLRIGSLALIHNTHLCPVHTV 480
 QY 481 PMDQFRNPQALHTANRPEDCVGBGLACHOLCARGHCWGPPTQCNVCSQFLRGQEC 540
 DB 481 PMDQFRNPQALHTANRPEDCVGBGLACHOLCARGHCWGPPTQCNVCSQFLRGQEC 540
 QY 541 VEECRVQLGLPRRYNARHCLPCHPCQPNQSVTCGPEADOCVACAHYKDPFCVARC 600
 DB 541 VEECRVQLGLPRRYNARHCLPCHPCQPNQSVTCGPEADOCVACAHYKDPFCVARC 600
 QY 601 PGGVKEPDLSTYMPKPEDEGACQPCPINTCSVDLDDKCPAPBASPLTISIAVVG 660
 DB 601 PGGVKEPDLSTYMPKPEDEGACQPCPINTCSVDLDDKCPAPBASPLTISIAVVG 660
 QY 661 ILVVVVLGVVFGILIRROQKIRKTYMRRLQSTELVEPLTPSGAMPNOAMRLKTEL 720
 DB 661 ILVVVVLGVVFGILIRROQKIRKTYMRRLQSTELVEPLTPSGAMPNOAMRLKTEL 720
 QY 721 RKRYKVGSGAFGVYVGIWIPDGENYKIPAIIVLRNTPKXNKETLDEAYVVAAGSP 780
 DB 721 RKRYKVGSGAFGVYVGIWIPDGENYKIPAIIVLRNTPKXNKETLDEAYVVAAGSP 780
 QY 781 YSRRLGICLTSTVQVLTQMPYGLLDHYRENRGRGSLDLNWCQIAKGSYLEBVR 840
 DB 781 YSRRLGICLTSTVQVLTQMPYGLLDHYRENRGRGSLDLNWCQIAKGSYLEBVR 840

QY 841 LVHRLAANRVTVKSNHVKITDPGLARLIDIDETEHADGKVPKIMMALISILRRRT 900
 DB 841 LVHRLAANRVTVKSNHVKITDPGLARLIDIDETEHADGKVPKIMMALISILRRRT 900
 QY 901 HOSDVSXGVTWELMTFGAKPYDGI PAREIPDLLEKGBLPQRPICCTIDVIMMKCM 960
 DB 901 HOSDVSXGVTWELMTFGAKPYDGI PAREIPDLLEKGBLPQRPICCTIDVIMMKCM 960
 QY 961 ISECRPRRELVSERSRMARDQRFVLIQNEIDLGPASPLDSTFYSLIEDDMGLVDA 1020
 DB 961 ISECRPRRELVSERSRMARDQRFVLIQNEIDLGPASPLDSTFYSLIEDDMGLVDA 1020
 QY 1021 EBYLVQOGFPCPDPAAGAGVHHRSSSTSGGGLTLGLSPSEEAAPSPLAPSEG 1080
 DB 1021 EBYLVQOGFPCPDPAAGAGVHHRSSSTSGGGLTLGLSPSEEAAPSPLAPSEG 1080
 QY 1081 AGSDVDFDGLGMAAGLQSLPTHDSPLQRYSEDPVPLPSESTDGVNAPLTCSPQPEYV 1140
 DB 1081 AGSDVDFDGLGMAAGLQSLPTHDSPLQRYSEDPVPLPSESTDGVNAPLTCSPQPEYV 1140
 QY 1141 NOPDVAPQPPSPREGPLPAAPAGATLERPKTSPGKNGVVDVAFGAVENPEYLTPQ 1200
 DB 1141 NOPDVAPQPPSPREGPLPAAPAGATLERPKTSPGKNGVVDVAFGAVENPEYLTPQ 1200
 QY 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPPERGAPSTFKGTPTAENPEYLGADV 1255
 DB 1201 GGAAPQHPHPAPSPAFDNLVYWDQDPPERGAPSTFKGTPTAENPEYLGADV 1255

RESULT 12
 AAE12130
 ID AAE12130 standard; Protein; 1255 AA.

AC AAE12130;
 XX 18-DEC-2001 (first entry)
 DT Human tyrosine kinase-type receptor, HER-2.
 DE
 XX
 XX Therapeutic compound; major histocompatibility complex; vaccine;
 KM antigenic peptide; MHC; immunoregulatory; immune response; HER-2;
 KM adoptive immunotherapy; anti-cancer; breast cancer antigen; APC;
 KM antigen presenting cell; human; tyrosine kinase-type receptor.
 XX
 XX Homo sapiens.

OS
 FH Key Location/Qualifiers
 FT Region 774..782
 FT /note="Antigenic epitope"

PN WO200168677-A2.

PD 20-SEP-2001.

PF 16-MAR-2001; 2001WO-US40328.

PR 16-MAR-2000; 2000US-0527487.

PA (GENZ) GENZYME CORP.

PI Nicolette CA;

DR WPI; 2001-616284/71.

DR N-PSDB; AAD19731.

PT Novel synthetic therapeutic compound for inducing immune response and
 PT for use in adoptive immunotherapy, has enhanced binding to major
 PT histocompatibility molecules and enhanced immunoregulatory properties

PS Claim 4; Page 63-67; 69pp; English.

CC The invention relates to synthetic therapeutic compounds (antigenic
 CC peptides) with enhanced binding to major histocompatibility complex
 CC (MHC) molecules and enhanced immunoregulatory properties relative
 CC to their natural counterparts. Compounds of the invention are useful
 CC for inducing an immune response in a subject and for use in adoptive
 CC immunotherapy. They are useful as components of anti-cancer vaccines
 CC and to expand immune effector cells that are specific for cancers
 CC characterized by expression of the breast cancer antigen, HER-2.
 CC Polynucleotides that encode peptides of the invention are useful as
 CC hybridization probes and as primers for the detection of genes of gene
 CC transcripts that are expressed in antigen presenting cells (APCs), to
 CC confirm transduction of polynucleotides into host cells. The present
 CC sequence is human tyrosine kinase-type receptor, HER-2. Compounds
 CC of the invention are designed based on the HER-2 antigenic peptide
 CC (774-782).

SO Sequence 1255 AA;

Query Match 99.9%; Score 6806; DB 22; Length 1255;

Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELAALCRKGLLALIPGAASVOVCTGDMKRLRASPETHLDMRLHYOGGVOVGNL 60
 DB 1 MELAALCRKGLLALIPGAASVOVCTGDMKRLRASPETHLDMRLHYOGGVOVGNL 60
 QY 61 ELTYLPTNASLFLQDIQEVGVYLAHNOVQVPLQRLIYVAGTOLFEDYNALAVLDNG 120
 DB 61 ELTYLPTNASLFLQDIQEVGVYLAHNOVQVPLQRLIYVAGTOLFEDYNALAVLDNG 120
 QY 121 DELNNTTPTVAGSPGGLRELQRLSTELLKGVLIQRPOLCYQDTILMKOIHKNQOLA 180
 DB 121 DELNNTTPTVAGSPGGLRELQRLSTELLKGVLIQRPOLCYQDTILMKOIHKNQOLA 180
 QY 181 LTLIDNBRACHPCSPMKSGSRCSSESDQSLTRTYCAGGACAKGGLPDDCCEOC 240
 DB 181 LTLIDNBRACHPCSPMKSGSRCSSESDQSLTRTYCAGGACAKGGLPDDCCEOC 240
 QY 241 AAGCTGPKSDCLACLFPHNSGICELHCPALVTYNTDFESMNPREGRYTFGASCVTACP 300
 DB 241 AAGCTGPKSDCLACLFPHNSGICELHCPALVTYNTDFESMNPREGRYTFGASCVTACP 300
 QY 301 YNVLSTDVSGCTLVCELAHQEVYABDGTORCEKSKPCARVCYGLGMEHLREYAVTSAN 360
 DB 301 YNVLSTDVSGCTLVCELAHQEVYABDGTORCEKSKPCARVCYGLGMEHLREYAVTSAN 360
 QY 361 IQEPACCKKIFGSLAFLPESFGDDPASNTAPLOPBOLOVETLEETGYLYISAMPDLP 420
 DB 361 IQEPACCKKIFGSLAFLPESFGDDPASNTAPLOPBOLOVETLEETGYLYISAMPDLP 420
 QY 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWGLSLRELSGLLIIHNTHLCFQHTV 480
 DB 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWGLSLRELSGLLIIHNTHLCFQHTV 480
 QY 481 PMDQLFRNPHQALLHTANRPEDECVBEGLAHQLCARGHMGSPPTQCVNCSQFLRGQBC 540
 DB 481 PMDQLFRNPHQALLHTANRPEDECVBEGLAHQLCARGHMGSPPTQCVNCSQFLRGQBC 540
 QY 541 VBECEVLOGLPRBYNARHCLPCHPECOPONGSVTCFGEADOCVACAHYKDPFCVARC 600
 DB 541 VBECEVLOGLPRBYNARHCLPCHPECOPONGSVTCFGEADOCVACAHYKDPFCVARC 600
 QY 601 PSGVKPDLSTYMPIMKPEDEGACQPCPINCHTSCVLDLDPGCAEORASPLTISIAYVG 660
 DB 601 PSGVKPDLSTYMPIMKPEDEGACQPCPINCHTSCVLDLDPGCAEORASPLTISIAYVG 660
 QY 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNQOMILKETEL 720
 DB 661 ILVVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNQOMILKETEL 720
 QY 721 RKRYVLGSGAGFTVYKGIWIPGSENVKIPYALIVLENTSPKANKELLDAYYVAGVSP 780
 DB 721 RKRYVLGSGAGFTVYKGIWIPGSENVKIPYALIVLENTSPKANKELLDAYYVAGVSP 780

QY 781 YSRRLGICLTSTVQVLTQMLPGYCLLDHYRENGRLGSODLLNMCQIAKMSYLEDVR 840
DB 781 YSRRLGICLTSTVQVLTQMLPGYCLLDHYRENGRLGSODLLNMCQIAKMSYLEDVR 840
QY 841 LVHRDLAARNVLYKSPNNHVKITDFGLARLLDIDETRYHADGKVPVKMMALBESTLRRT 900
DB 841 LVHRDLAARNVLYKSPNNHVKITDFGLARLLDIDETRYHADGKVPVKMMALBESTLRRT 900
QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPPOPCTIDVYIMVCKM 960
DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPPOPCTIDVYIMVCKM 960
QY 961 IDSECRFRFRELVESESRMARDPFRVYIQNEDLGPSPLDSTYRSLLEDDMDGLVDA 1020
DB 961 IDSECRFRFRELVESESRMARDPFRVYIQNEDLGPSPLDSTYRSLLEDDMDGLVDA 1020
QY 1021 EETLVPOQGFPCPDPAAGAGMVRHRRSSSTRSGGDLTLGLEPSREBA PRSPAPSEG 1080
DB 1021 EETLVPOQGFPCPDPAAGAGMVRHRRSSSTRSGGDLTLGLEPSREBA PRSPAPSEG 1080
QY 1081 AGSDVPEGDGLMGAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGYAAPLTCSPQPEYV 1140
DB 1081 AGSDVPEGDGLMGAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGYAAPLTCSPQPEYV 1140
QY 1141 NQSDVPRPQPSRPRGFLPAARPAATLERAKTSLPGKNGVVKOVFAFGAVENPEYLTPO 1200
DB 1141 NQSDVPRPQPSRPRGFLPAARPAATLERAKTSLPGKNGVVKOVFAFGAVENPEYLTPO 1200
QY 1201 GGAAPQHPHPAPSPAPDNLTYMDQDPPRGPAPSTFKGTPTANPEYLGADVAV 1255
DB 1201 GGAAPQHPHPAPSPAPDNLTYMDQDPPRGPAPSTFKGTPTANPEYLGADVAV 1255

RESULT 13
AAB60167
ID AAB60167 standard; Protein; 1255 AA.
AC AAB60167;
XX
XX
XX 03-APR-2001 (first entry)
DT
XX
XX
DE HER2 transgene plasmid construct encoded protein.
XX
XX
KW Human; HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;
XX
XX
OS antibody.
XX
OS Homo sapiens.
XX
XX
PN MO200100244-A2.
XX
PD 04-JAN-2001.
XX
XX
PF 23-JUN-2000; 2000MO-US17229.
XX
XX
PR 25-JUN-1999; 99US-0141316.
XX
XX
PR 16-MAR-2000; 2000US-0189844.
XX
XX
PA (GETH) GENENTECH INC.
XX
XX
PI Erickson S, Schwall R;
XX
XX
DR MPI; 2001-061962/07.
XX
XX
DR N-PSDB; AAF24297.
XX
XX
PT Treating tumors, particularly breast cancers, which overexpress an ErbB
XX
XX
PT receptor and does not respond to an anti-ErbB antibody, comprises
XX
XX
PT conjugating the antibody to a maytansinoid -
XX
XX
XX Example 3; Fig 4; 92pp; English.
XX
XX
CC The present invention provides a method of treating cancer by

CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In
CC particular, the antibody is directed against ErbB2 (also known as HER2
CC and p185neu). The method is particularly useful in the treatment of
CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,
CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.

SO Sequence 1255 AA;
Query Match 99.9%; Score 6806; DB 22; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 METALACRMGLLALLPPGAASSTOVCTGDMKRLPASPETHLMDLHLYOGCOVOGNTL 60
DB 1 METALACRMGLLALLPPGAASSTOVCTGDMKRLPASPETHLMDLHLYOGCOVOGNTL 60
QY 61 EETLVPTNASTSLFLODIQEVQGVVLLAHNOVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
DB 61 EETLVPTNASTSLFLODIQEVQGVVLLAHNOVROVPLQRLIRVGTQLFEDNYALAVLDNG 120
QY 121 DFLANTTPTVYASPGGLRELQRLSTELLKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
DB 121 DFLANTTPTVYASPGGLRELQRLSTELLKGVLIQNPOLCYODTILMKDIFHKNNOLA 180
QY 181 LHLIDTNRSPACHPSPCKSGRCKWBSSEDCSLTRTVACAGCARCKGPLPTDCHEQC 240
DB 181 LHLIDTNRSPACHPSPCKSGRCKWBSSEDCSLTRTVACAGCARCKGPLPTDCHEQC 240
QY 241 ANCTGPKHSDCLACHFNHSGICELHCPALVYNTDTPFSMPNBERYTFGASCVTACP 300
DB 241 ANCTGPKHSDCLACHFNHSGICELHCPALVYNTDTPFSMPNBERYTFGASCVTACP 300
QY 301 YNLTSTDVSGCTLVCPDLAHNOEVTAEQTORCEKSKRCARVCYGLGMEHLREVAVTSAN 360
DB 301 YNLTSTDVSGCTLVCPDLAHNOEVTAEQTORCEKSKRCARVCYGLGMEHLREVAVTSAN 360
QY 361 IQEPACCKKIFGSLATLPESFPDGPASNTAPLOPEOLQVETLEITGYLYISAMPDLP 420
DB 361 IQEPACCKKIFGSLATLPESFPDGPASNTAPLOPEOLQVETLEITGYLYISAMPDLP 420
QY 421 DISVPONLOVYIRIRILHNGAVSLTLQIGISWGLSLRLSGSLAIHNTNHCYFHTV 480
DB 421 DISVPONLOVYIRIRILHNGAVSLTLQIGISWGLSLRLSGSLAIHNTNHCYFHTV 480
QY 481 PMDQLFRNPHQALHTANPEDECVSGGLACHOLCARHGMCGPPTOCVACSOFLRQEC 540
DB 481 PMDQLFRNPHQALHTANPEDECVSGGLACHOLCARHGMCGPPTOCVACSOFLRQEC 540
QY 541 VEECRVLOGLPRRYVNAHCLPCHPECOPONGSVTCGPADOCVACAHYKDPFCVARC 600
DB 541 VEECRVLOGLPRRYVNAHCLPCHPECOPONGSVTCGPADOCVACAHYKDPFCVARC 600
QY 601 PSGVXPDLSTMPIMKPPDEEGACOPCPINCTHSCVDLDDGCPAEORASPLTISIVVG 660
DB 601 PSGVXPDLSTMPIMKPPDEEGACOPCPINCTHSCVDLDDGCPAEORASPLTISIVVG 660
QY 661 ILLVVVLGVVFGILLIRROOKIRKTYMRLLQETELVEPLTPSGAMENQAMRLKETEL 720
DB 661 ILLVVVLGVVFGILLIRROOKIRKTYMRLLQETELVEPLTPSGAMENQAMRLKETEL 720
QY 721 RRYKTVLGSAGFTVYKGIWI PDGENYKIPVAILVLRBENTS PKANKETLDEAYVAGVSP 780
DB 721 RRYKTVLGSAGFTVYKGIWI PDGENYKIPVAILVLRBENTS PKANKETLDEAYVAGVSP 780
QY 781 YSRRLGICLTSTVQVLTQMLPGYCLLDHYRENGRLGSODLLNMCQIAKMSYLEDVR 840
DB 781 YSRRLGICLTSTVQVLTQMLPGYCLLDHYRENGRLGSODLLNMCQIAKMSYLEDVR 840
QY 841 LVHRDLAARNVLYKSPNNHVKITDFGLARLLDIDETRYHADGKVPVKMMALBESTLRRT 900
DB 841 LVHRDLAARNVLYKSPNNHVKITDFGLARLLDIDETRYHADGKVPVKMMALBESTLRRT 900
QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPPOPCTIDVYIMVCKM 960

```

DB 901 HOSDWSYGVTVWELMTFGAKPYDGI.PABEIPDLLEKGERLPQPCTIVIMVACMM 960
QY 961 IDSECRPRRELVSSESRMARDFORFVVIQNEBIDGPASPLDSFPPYRSLBDDMGDLVDA 1020
DB 961 IDSECRPRRELVSSESRMARDFORFVVIQNEBIDGPASPLDSFPPYRSLBDDMGDLVDA 1020
QY 1021 BEYLVPQQGFPCDPAPAGAGMVHRRHSSTRSGGDLTLGLEPSEEA PRSPLAPSEG 1080
DB 1021 BEYLVPQQGFPCDPAPAGAGMVHRRHSSTRSGGDLTLGLEPSEEA PRSPLAPSEG 1080
QY 1081 AGSDVDGDLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVAVLTCSPQPEYV 1140
DB 1081 AGSDVDGDLGMAAGLQSLPTHDSPLQRYSEDPVPLPSETDGVAVLTCSPQPEYV 1140
QY 1141 NOPDVAPQPPSPREGPLPAARPAATLERPKTSLPGKGVVXDPVAFAGAVENPEYLPQ 1200
DB 1141 NOPDVAPQPPSPREGPLPAARPAATLERPKTSLPGKGVVXDPVAFAGAVENPEYLPQ 1200
QY 1201 GGAAPQHPPEPAPSPAFDNLVYWDOPPERGAPSTFKGPTTANPEYLGLDVPV 1255
DB 1201 GGAAPQHPPEPAPSPAFDNLVYWDOPPERGAPSTFKGPTTANPEYLGLDVPV 1255

RESULT 14
AAE26349 standard; Protein; 1255 AA.
XX AC AAE26349;
XX DT 13-DEC-2002 (first entry)
XX DE Human HER-2 protein.
XX KM Transgenic animal; transgenic; mammary gland cell; HER2; tumour;
XX KM cancer; therapy; apoptosis; cytostatic; human.
XX OS Homo sapiens.
XX PN US2002035736-A1.
XX PD 21-MAR-2002.
XX PF 16-MAR-2001; 2001US-0811115.
XX PR 16-MAR-2000; 2000US-189844P.
XX PA (ERIC/) ERICKSON S.
XX PA (KING/) KING K.
XX PA (SCHW/) SCHWALL R.
XX PI Erickson S, King K, Schwall R;
XX DR WPI; 2002-401155/43.
XX DR N-PSDB; AAD43934, AAD43935.
XX PT New transgenic non-human mammal that produces detectable levels of a
XX PT native human HER2 protein in its mammary gland cells, useful as tumor
XX PT models for testing HER2-directed cancer therapies, and for identifying
XX PT anticancer agents
XX PS Example 2; Page 26-29; 83pp; English.
XX XX The invention relates to a transgenic non-human mammal that produces in
XX XX its mammary gland cells detectable levels of a native human HER2 protein
XX XX or its fragment. The transgenic animals are useful as tumour models for
XX XX testing HER2-directed cancer therapies, and for identifying anticancer
XX XX agents. The animals may also be used as source of cells which can be
XX XX immortalised in culture, in screening for compounds that have potential
XX XX as prophylactic or therapeutic treatments of diseases or disorders
XX XX involving expression of HER2. The anti-cancer molecules are useful for
XX XX inducing apoptosis or cell death of cancer cells. The present sequence
XX XX is human HER-2 protein.

```

```

XX SQ Sequence 1255 AA;
QY 1 MEALALCRGGLLALALPPGASASTQVCTGDMKRLPASPEETHLDMRLHYOGCQVQGNL 60
DB 1 MEALALCRGGLLALALPPGASASTQVCTGDMKRLPASPEETHLDMRLHYOGCQVQGNL 60
QY 61 EUTYLPFNASLSFLQDIQEVQGVLLIAHQVROVPLQRLIRVGTQLFEDNVALAVLDNG 120
DB 61 EUTYLPFNASLSFLQDIQEVQGVLLIAHQVROVPLQRLIRVGTQLFEDNVALAVLDNG 120
QY 121 DPLNNTPTVAGSPGRLRELQSLRLLEIKGGLIQNPOLCYQDPTLMDIFHKNNOLA 180
DB 121 DPLNNTPTVAGSPGRLRELQSLRLLEIKGGLIQNPOLCYQDPTLMDIFHKNNOLA 180
QY 181 LTLIDNRRARACHPCSPMKGSRGMSSESDCSLRTVCAAGCARCKGPLPTDCHEQC 240
DB 181 LTLIDNRRARACHPCSPMKGSRGMSSESDCSLRTVCAAGCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPDEGRYTFGASCVTACP 300
DB 241 AAGCTGPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPDEGRYTFGASCVTACP 300
QY 301 YNLTSTDVSCCTVCPDINQOVTAEQGTORCEKSRPCARVCYGLGMEHLREVAATVSAN 360
DB 301 YNLTSTDVSCCTVCPDINQOVTAEQGTORCEKSRPCARVCYGLGMEHLREVAATVSAN 360
QY 361 IDEFACCKKIFGSLAPLPESPQDPAANTAPLOPELOVETLEITGYLISAMPDPL 420
DB 361 IDEFACCKKIFGSLAPLPESPQDPAANTAPLOPELOVETLEITGYLISAMPDPL 420
QY 421 DLSVFQNLQVIRGRILHANGAYSITLQGLGISWGLRLSRLSGGLALIHNNTHLCFVHTV 480
DB 421 DLSVFQNLQVIRGRILHANGAYSITLQGLGISWGLRLSRLSGGLALIHNNTHLCFVHTV 480
QY 481 PMDQLFRNPHOALHFRANRPEDECVESGLACHQLCARGHCKGFGPPQCVNCSQFLRQEC 540
DB 481 PMDQLFRNPHOALHFRANRPEDECVESGLACHQLCARGHCKGFGPPQCVNCSQFLRQEC 540
QY 541 VEECRVLOGLPREYVVARCLPCHPECOPONGSVTFEGPEADQVCAHAKDPFCVAC 600
DB 541 VEECRVLOGLPREYVVARCLPCHPECOPONGSVTFEGPEADQVCAHAKDPFCVAC 600
QY 601 PGGVKEDLSYMPIWKEPDEEGACQPCPINCTHSCVDLDKGCPRABORASPLTISAVVG 660
DB 601 PGGVKEDLSYMPIWKEPDEEGACQPCPINCTHSCVDLDKGCPRABORASPLTISAVVG 660
QY 661 ILLVVLGVVFGILLKRRQOKTRKTYMRLLQETLVEPLTPSGAMPNOQMRLKETEL 720
DB 661 ILLVVLGVVFGILLKRRQOKTRKTYMRLLQETLVEPLTPSGAMPNOQMRLKETEL 720
QY 721 RMYKVLGSGAPGVYVGIWIPGSENVKIPVATVLEBNTSPKANKELIDEAYVAGVGP 780
DB 721 RMYKVLGSGAPGVYVGIWIPGSENVKIPVATVLEBNTSPKANKELIDEAYVAGVGP 780
QY 781 YVSRLLGICLTSTVQLVTLQMPYGLLDHVRBNRGLSGODLNMCMQIAXKGSYLEDDR 840
DB 781 YVSRLLGICLTSTVQLVTLQMPYGLLDHVRBNRGLSGODLNMCMQIAXKGSYLEDDR 840
QY 841 LVHRDLAANVLYKSENNHYKTDFGLARLIDETRYHADGKVPYKMALESILRRRFT 900
DB 841 LVHRDLAANVLYKSENNHYKTDFGLARLIDETRYHADGKVPYKMALESILRRRFT 900
QY 901 HOSDWSYGVTVWELMTFGAKPYDGI.PABEIPDLLEKGERLPQPCTIVIMVACMM 960
DB 901 HOSDWSYGVTVWELMTFGAKPYDGI.PABEIPDLLEKGERLPQPCTIVIMVACMM 960
QY 961 IDSECRPRRELVSSESRMARDFORFVVIQNEBIDGPASPLDSFPPYRSLBDDMGDLVDA 1020
DB 961 IDSECRPRRELVSSESRMARDFORFVVIQNEBIDGPASPLDSFPPYRSLBDDMGDLVDA 1020

```

Db 961 IDSECRPRRELVSERSRMARDPQRVYVIONEDLGPAASPLDSTFYNSLLEDDMDGLVDA 1020
 Qy 1021 EBYLVPOQGFPCDPAPAGAGVHHRRSSSTSSGGGDLTLGLEPSEEARPSPLAPSEG 1080
 Db 1021 EBYLVPOQGFPCDPAPAGAGVHHRRSSSTSSGGGDLTLGLEPSEEARPSPLAPSEG 1080
 Qy 1081 AGSDVDVDGDLGMAAGLGLSLPTHDPSPLORYSEDPVPLPSTEDGVVALTSGPQPEYV 1140
 Db 1081 AGSDVDVDGDLGMAAGLGLSLPTHDPSPLORYSEDPVPLPSTEDGVVALTSGPQPEYV 1140
 Qy 1141 NOPDVAPPPSPREGPLPAARPAATLERPKTSPGKGVVKKVPAFGAVENPEYLTPO 1200
 Db 1141 NOPDVAPPPSPREGPLPAARPAATLERPKTSPGKGVVKKVPAFGAVENPEYLTPO 1200
 Qy 1201 GGAAPOPHPPPAPSPAFDNLVYWDQDPERFGAPSTFKGTPTAENPEYLGLDVAV 1255
 Db 1201 GGAAPOPHPPPAPSPAFDNLVYWDQDPERFGAPSTFKGTPTAENPEYLGLDVAV 1255
 RESULT 15
 AAE26366
 ID AAE26366 standard; Protein; 1255 AA.
 AC AAE26366;
 XX 13-DEC-2002 (first entry)
 DE Human Her2 antigen.
 XX Human: Immune response; T-helper cell epitope; chitosan; CTL response;
 KM vaccine; prostate cancer; breast cancer; Her2 antigen; cytostatic;
 KM immunostimulant.
 XX Homo sapiens.
 OS
 FH Key location/Qualifiers
 FT Peptide 1..23
 FT Protein /label= signal_peptide
 FT 24..1255
 /note= "Mature human Her2 antigen"
 PN MO200234287-A2.
 XX 02-MAY-2002.
 XX 26-OCT-2001; 2001WO-DK00705.
 XX 27-OCT-2000; 2000DK-0001606.
 PR 03-NOV-2000; 2000US-245166P.
 PR 18-JUN-2001; 2001DK-0000936.
 XX (PHAR-) PHARMEXA AS.
 PA Beier AM, Gautam A, Mouritsen S;
 PI WPI; 2002-463339/49.
 DR N-PSDB; AAD43986.
 XX
 PT Inducing or enhancing an immune response against an antigen,
 PT particularly cytotoxic T-lymphocyte responses, for treating or
 PT ameliorating prostate or breast cancer, comprises administering the
 PT antigen formulated with chitosan
 XX
 PS Disclosure; Page 91-95; 97pp; English.
 XX
 CC The invention relates to a method for inducing or enhancing an immune
 CC response against a polypeptide antigen in an animal, including human.
 CC The method comprises administering the polypeptide antigen or at least
 CC one variant which includes at least one first T-helper cell epitope that
 CC is foreign to the animal (foreign T_H epitope) and is formulated with
 CC chitosan. The polypeptide antigen is weakly immunogenic or non-
 CC immunogenic. The invention is used as vaccine. The chitosan and
 CC polypeptide antigen or its variant are useful in the preparation of an

CC immunogenic composition for inducing or enhancing an immune response,
 CC particularly CTL response, against the polypeptide or protein antigen.
 CC The method for inducing or enhancing an immune response is useful in
 CC treating or ameliorating cancer, e.g. prostate or breast cancer. The
 CC present sequence is human Her2 antigen.
 XX
 SQ Sequence 1255 AA;
 Query Match 99.9%; Score 6806; DB 23; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MELAALCRMGILLALIPGAASTOVCTGDMKRLPASPETHIDMLRHLYOGGCOVQGNL 60
 Db 1 MELAALCRMGILLALIPGAASTOVCTGDMKRLPASPETHIDMLRHLYOGGCOVQGNL 60
 Qy 61 ELYLVPTNLSLSTLOIOEVQGVYLLAHNOVRQVPLQRLRIYVGTOLEFEDNYALAVLDNG 120
 Db 61 ELYLVPTNLSLSTLOIOEVQGVYLLAHNOVRQVPLQRLRIYVGTOLEFEDNYALAVLDNG 120
 Qy 121 DELNNTPTVYTGASPGGLRELQRLSTLEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
 Db 121 DELNNTPTVYTGASPGGLRELQRLSTLEILKGVLIQRPOLCYQDTILMKDIFHKNNOLA 180
 Qy 181 LTLIDPFRSRACPCSPMCKGSRQWGSSESDCSLRTVCAGGACRCCKPLPTDCHEOC 240
 Db 181 LTLIDPFRSRACPCSPMCKGSRQWGSSESDCSLRTVCAGGACRCCKPLPTDCHEOC 240
 Qy 241 AAGCTGPKHSIDCLACHFNHSGICELHCPALVYNTDTFESMNPREGYTFGASCVTACP 300
 Db 241 AAGCTGPKHSIDCLACHFNHSGICELHCPALVYNTDTFESMNPREGYTFGASCVTACP 300
 Qy 301 YNTLSTDVSGCTLVCPLANOEVTABEDTQRCCKSPCARVCYGLGMEHLREYRAVTSAN 360
 Db 301 YNTLSTDVSGCTLVCPLANOEVTABEDTQRCCKSPCARVCYGLGMEHLREYRAVTSAN 360
 Qy 361 IOEFACCKKIFGSLAFLPSPFDGDPASNTAPLOPELOVETLEBETGYIYSAMPDLP 420
 Db 361 IOEFACCKKIFGSLAFLPSPFDGDPASNTAPLOPELOVETLEBETGYIYSAMPDLP 420
 Qy 421 DISVFQNLQVIRIRILHNGAYSLTLQGLISWGLSLRLSGSLALIHNTHLCFVHTV 480
 Db 421 DISVFQNLQVIRIRILHNGAYSLTLQGLISWGLSLRLSGSLALIHNTHLCFVHTV 480
 Qy 481 PMDQLFRNPHQALLHTANRPEDECVSGLACHQLCARGHCWGPGPTCVNCSQPLRQEC 540
 Db 481 PMDQLFRNPHQALLHTANRPEDECVSGLACHQLCARGHCWGPGPTCVNCSQPLRQEC 540
 Qy 541 VBECRVLQGLPREYVARHCLPCHPSCOPONGSVTFGEPADOCVACAHYKDPFCVARC 600
 Db 541 VBECRVLQGLPREYVARHCLPCHPSCOPONGSVTFGEPADOCVACAHYKDPFCVARC 600
 Qy 601 PSCVKEPDLSTMPYKPEDEGACQPCPINCTHSCVDLDDKGCFAEGRASPLTISAVVG 660
 Db 601 PSCVKEPDLSTMPYKPEDEGACQPCPINCTHSCVDLDDKGCFAEGRASPLTISAVVG 660
 Qy 661 ILVAVVVGIVFGLIKRRQCKIRKTYMRRLQSTELVEPLTPSGAMNQAOMRLKTEL 720
 Db 661 ILVAVVVGIVFGLIKRRQCKIRKTYMRRLQSTELVEPLTPSGAMNQAOMRLKTEL 720
 Qy 721 RKRYKVGSGAFGVYGIWIPDEGNVYKIPVAKVLRNTPSKNKEILDEAYVYAGVGP 780
 Db 721 RKRYKVGSGAFGVYGIWIPDEGNVYKIPVAKVLRNTPSKNKEILDEAYVYAGVGP 780
 Qy 781 YSRRLGICLTSTVOLVTLMPYGCILLDHVRENKRGSLDLMNCQIAKMSYLEDVR 840
 Db 781 YSRRLGICLTSTVOLVTLMPYGCILLDHVRENKRGSLDLMNCQIAKMSYLEDVR 840
 Qy 841 LVHRDLAARNVLYKSPNHVKTTPGRLALDIDETEHAGKVPKIMMALESILRRFT 900
 Db 841 LVHRDLAARNVLYKSPNHVKTTPGRLALDIDETEHAGKVPKIMMALESILRRFT 900
 Qy 901 HQSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGBRLPOPICTIDVYIMVNCWM 960

Db	901	HQSDVMSYGVTTWELMTFGAKFYDGIIPAREIPDLLEKGERLPDPPICTIDVTMTWKMM	960
Qy	961	IDSECRPRFRELVSFSSRMARDPQRFVVIQNEIDGPASPLDSTFYRSLLBDDMDGLVDA	1020
Db	961	IDSECRPRFRELVSFSSRMARDPQRFVVIQNEIDGPASPLDSTFYRSLLBDDMDGLVDA	1020
Qy	1021	EETLVPOQGFPCDDPAPGAGMWHHRSSSTRSGGDLTLGLEPSEEA PRSPLAPSEG	1080
Db	1021	EETLVPOQGFPCDDPAPGAGMWHHRSSSTRSGGDLTLGLEPSEEA PRSPLAPSEG	1080
Qy	1081	AGSDVDFGDLGMAKGLQSLPTHTDPSPLORYSEDPVTPLPSHTDGYVAPLTCSPQPEYV	1140
Db	1081	AGSDVDFGDLGMAKGLQSLPTHTDPSPLORYSEDPVTPLPSHTDGYVAPLTCSPQPEYV	1140
Qy	1141	NQPDVAPQPPSPREGPLPAARPAATLERPKTLPKNGVVKOVFAFGAVENPEYLTPO	1200
Db	1141	NQPDVAPQPPSPREGPLPAARPAATLERPKTLPKNGVVKOVFAFGAVENPEYLTPO	1200
Qy	1201	GGAAPQPHPPAPSPAFDNLTYMDQDPPERGA PPSFTKGTPTAENPEYLGLDVAV	1255
Db	1201	GGAAPQPHPPAPSPAFDNLTYMDQDPPERGA PPSFTKGTPTAENPEYLGLDVAV	1255

Search completed: February 9, 2004, 16:15:14
 Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: February 9, 2004, 16:13:07; Search time 27 Seconds

(without alignments)
4470.063 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815
Sequence: 1 MEALALCRWGLLLALPEGA.....TFKGTPTAENPEYLGLDPV 1255Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

PIR 76: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6815	100.0	1255	1 A24571	protein-tyrosine k
2	5997	88.0	1260	1 TVRTMU	protein-tyrosine k
3	5993.5	87.9	1254	2 I481C1	P-185 precursor -
4	3169	46.5	1210	1 GQHUR	epidermal growth f
5	3145	46.1	1210	2 A531B3	epidermal growth f
6	3122.5	45.8	1223	1 TVCHLV	epidermal growth f
7	3004.5	44.1	1308	2 A47253	epidermal growth f
8	2708	39.7	1166	1 S06142	protein-tyrosine k
9	2432.5	35.7	1342	2 A36223	kinase-related tra
10	2347.5	34.4	1339	1 UC43B7	epidermal growth f
11	1765.5	25.9	698	1 TVFVIV	protein-tyrosine k
12	1702	25.0	604	1 TVYUH	epidermal growth f
13	1653	24.3	1330	1 GQFPE	epidermal growth f
14	1645	24.1	544	2 S35745	protein-tyrosine k
15	1638	24.0	545	2 S00727	kinase-related tra
16	1621	23.8	540	2 B44776	protein-tyrosine k
17	1619	23.5	540	1 TVFVVB	protein-tyrosine k
18	1536	22.5	644	2 A36355	epidermal growth f
19	1301	19.1	1323	2 B88257	protein let-23 [lm
20	1301	19.1	1374	2 S70712	protein-tyrosine k
21	1211	17.8	1369	2 S70713	protein-tyrosine k
22	1177	17.3	1717	1 A45558	epidermal growth f
23	1155	16.9	527	2 A42031	epidermal growth f
24	997.5	14.6	843	2 A27131	epidermal growth f
25	814.5	12.0	346	2 S13807	protein-tyrosine k
26	757	11.1	311	2 S13808	protein-tyrosine k
27	736	10.8	1363	2 T43220	insulin-like growt
28	717	10.5	1382	1 INHUR	insulin receptor p
29	710	10.4	1383	2 A36080	insulin receptor p

30	709.5	10.4	1372	2 A34157	insulin receptor p
31	703.5	10.3	1607	2 T43212	insulin-like growt
32	693.5	10.2	1300	2 A36502	insulin receptor-r
33	682.5	10.0	1477	2 T18534	protein-tyrosine k
34	681	10.0	1268	2 B36502	insulin receptor-r
35	651	9.6	1367	1 IGHUR1	insulin-like growt
36	642	9.4	1371	2 A33837	insulin-like growt
37	627	9.2	1390	2 T30346	insulin receptor -
38	624.5	9.2	2148	1 A56081	insulin receptor -
39	622	9.1	2101	2 S57245	insulin receptor (
40	607	8.9	987	2 A54052	protein-tyrosine k
41	591	8.7	1091	2 S33596	protein-tyrosine k
42	590.5	8.7	952	2 I50612	protein-tyrosine k
43	589.5	8.7	977	2 S49004	tyrosine kinase Mp
44	587	8.6	1114	1 S05582	protein-tyrosine k
45	585	8.6	987	2 I48652	mouse developmenta

ALIGNMENTS

RESULT 1

A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N/Alternate names: C-erb-B-2 protein precursor; Kinase-related transforming protein erb
C/Species: Homo sapiens (man)
C/Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999
C/Accession: A24571, A25491, A44188, B44188, I59509, I57622
R/Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nature 319, 230-234, 1986
A/Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth
A/Reference number: A24571, MUID:66118663, PMID:3303577
A/Accession: A24571
A/Molecule type: mRNA
A/Residues: 1-1255 <YAM>
A/Cross-references: GB:X03363, NID:G31197, PIDN:CAA27060.1, PID:G31198
R/Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A/Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epider
A/Reference number: A25491, MUID:66016723, PMID:2995967
A/Accession: A25491
A/Molecule type: DNA
A/Residues: 737-1031 <SEM>
A/Molecule type: DNA
A/Cross-references: GB:M11767, NID:G182163, PIDN:AAA35808.1, PID:G553282
R/Cousens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, J.
Science 230, 1132-1139, 1985
A/Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromo
A/Reference number: A44188, MUID:86070181, PMID:2999974
A/Accession: A44188
A/Molecule type: DNA
A/Residues: 740-910 <COU1>
A/Cross-references: GB:M12036, NID:G183988, PIDN:AAA35978.1, PID:G183989
A/Accession: B44188
A/Molecule type: mRNA
A/Residues: 1-517, 'RALU', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A/Cross-references: GB:M11730, NID:G183986
R/King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 228, 974-976, 1985
A/Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A/Reference number: I59509, MUID:85272597, PMID:2992089
A/Accession: I59509
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 832-909 <REX>
A/Cross-references: GB:L29395, NID:G459807, PIDN:AAA35809.1, PID:G459808
R/Tal, M.; King, C.R.; Kraus, M.H.; Ulrich, A.; Schleisinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A/Title: Human HBR2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A/Reference number: I57622, MUID:87268898, PMID:3039351
A/Accession: I57622
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-191 <TML>

A/Cross-references: GB:M16792; NID:g183983; PID:AAA58637.1; PID:g553332
 C/Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C/Genetics:
 A/Gene: GDB:ERBB2; NGL; NEU; HER-2
 A/Cross-references: GDB:120613; OMIM:164870
 A/Map position: 17q21.1-17q21.1
 A/Introns: 25/1; 75/3; 147/1; 883/3
 A/Note: the list of introns is incomplete
 C/Function:
 A/Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphinase
 F/1-21/Domain: signal sequence #status predicted <SIG>
 F/22-125/Produce: protein-tyrosine kinase erbB2 #status predicted <MAT>
 F/22-653/Domain: extracellular #status predicted <EXT>
 F/70-304/Domain: EGF receptor extracellular domain repeat <EB1>
 F/395-605/Domain: EGF receptor extracellular domain repeat <EB2>
 F/554-675/Domain: transmembrane #status predicted <TM>
 F/676-1255/Domain: intracellular #status predicted <INT>
 F/718-983/Domain: protein kinase homology <KIN>
 F/726-734/Region: protein kinase ATP-binding motif
 F/68.124.187.259.530.571.629/Binding site: carbohydrate (asn) (covalent) #status predict
 F/686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F/753/Active site: Lys #status predicted
 F/1139.1221.1222.1246/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 100.0%; Score 6815; DB 1; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1.1e-270;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEALALCRWGLTALLPRGAATVCTGTDMKRLPASPTHLDMLRLHYOGCCVVOGNTL 60
 Db 1 MEALALCRWGLTALLPRGAATVCTGTDMKRLPASPTHLDMLRLHYOGCCVVOGNTL 60
 Qy 61 ELTYLPNALSLSQDIQEVGYTLIAHQVQVPLQRLRVRTQVLFEDVYALAVNDNG 120
 Db 61 ELTYLPNALSLSQDIQEVGYTLIAHQVQVPLQRLRVRTQVLFEDVYALAVNDNG 120
 Qy 121 DPLNNTPTVPGASGGLRELOLRSLTEILKGVLIQNRPOLCYODTILMKDIFRKNOLA 180
 Db 121 DPLNNTPTVPGASGGLRELOLRSLTEILKGVLIQNRPOLCYODTILMKDIFRKNOLA 180
 Qy 181 LTIIDNRSRACHPCSPWCKSGSRGSSBDQSLTRTVACGACRCKGRLPTDCHEQC 240
 Db 181 LTIIDNRSRACHPCSPWCKSGSRGSSBDQSLTRTVACGACRCKGRLPTDCHEQC 240
 Qy 241 AAGCTGRKSHDCLAHLPHNSGTEHLNCPALVYNTDTFESMPNPEGRYTFGASCVTAAC 300
 Db 241 AAGCTGRKSHDCLAHLPHNSGTEHLNCPALVYNTDTFESMPNPEGRYTFGASCVTAAC 300
 Qy 301 YNYLSTVGSCTVLCPLHNOEVTAEQTCRCKSKPCARCYGLGMEHLREVAATYSAN 360
 Db 301 YNYLSTVGSCTVLCPLHNOEVTAEQTCRCKSKPCARCYGLGMEHLREVAATYSAN 360
 Qy 361 IOEPAGCKKIFGSLAFLPESFDGPASNTAFLQEPOLQVFETLEITGYLYISAMPDLP 420
 Db 361 IOEPAGCKKIFGSLAFLPESFDGPASNTAFLQEPOLQVFETLEITGYLYISAMPDLP 420
 Qy 421 DLSVFOULQVIRGAILNNGAYSLTLOGLIGISWGLRSLREISGSLALHNHTHLCPHYTV 480
 Db 421 DLSVFOULQVIRGAILNNGAYSLTLOGLIGISWGLRSLREISGSLALHNHTHLCPHYTV 480
 Qy 481 PMDLPFNPHQALHTNARPEDECVGEGTACHOLCARHCHGPGPTQCVNCSQFLRGEC 540
 Db 481 PMDLPFNPHQALHTNARPEDECVGEGTACHOLCARHCHGPGPTQCVNCSQFLRGEC 540
 Qy 541 VEEGRVLOGLPREVYNAHCLPCHPECCOPONGSYTCFPEADQCVACAHYDPPFCVAC 600
 Db 541 VEEGRVLOGLPREVYNAHCLPCHPECCOPONGSYTCFPEADQCVACAHYDPPFCVAC 600
 Qy 601 PGVGPDLSTYPIKPPDEBACOPPCINCTHSCVDLDDKCPABQASPLTSTISAVG 660
 Db 601 PGVGPDLSTYPIKPPDEBACOPPCINCTHSCVDLDDKCPABQASPLTSTISAVG 660

Db 601 PGVGPDLSTYPIKPPDEBACOPPCINCTHSCVDLDDKCPABQASPLTSTISAVG 660
 Qy 661 ILVVVLGVVFGILLIKRRQOKIRKTYMRLLQSTELVEPLTPSGAMENQOMRILKETEL 720
 Db 661 ILVVVLGVVFGILLIKRRQOKIRKTYMRLLQSTELVEPLTPSGAMENQOMRILKETEL 720
 Qy 721 RAKVYUGSAGFVYVYGIWTPDEENKIPVAIKVLENTSPKANKELDEAYVNAAGVSP 780
 Db 721 RAKVYUGSAGFVYVYGIWTPDEENKIPVAIKVLENTSPKANKELDEAYVNAAGVSP 780
 Qy 781 YVRLRGLICTSTVQVLTQMPYGLLDHYENRGRGLSODLNMCMQIAKMSYLEVDV 840
 Db 781 YVRLRGLICTSTVQVLTQMPYGLLDHYENRGRGLSODLNMCMQIAKMSYLEVDV 840
 Qy 841 LVPRDLAARNVLYKSPNHVKTDFGLARLLDIDETEHADGKVPIMMALESLRRFT 900
 Db 841 LVPRDLAARNVLYKSPNHVKTDFGLARLLDIDETEHADGKVPIMMALESLRRFT 900
 Qy 901 HQSDVMSYGVTTWELMTFGAKPYDGIIPAREIPDLLEGERLPORCTIVYMTMVKCM 960
 Db 901 HQSDVMSYGVTTWELMTFGAKPYDGIIPAREIPDLLEGERLPORCTIVYMTMVKCM 960
 Qy 961 IDSECRPRFELVSEFSRMAKDPORFVYIQNBDLGPASPLDSFYSRLLEDMDGLVDA 1020
 Db 961 IDSECRPRFELVSEFSRMAKDPORFVYIQNBDLGPASPLDSFYSRLLEDMDGLVDA 1020
 Qy 1021 EETLVDPQGFCEPDPAFAGAGVHHRRSSSTRSGGDLTLGLEPSEEARSPPLAPSEG 1080
 Db 1021 EETLVDPQGFCEPDPAFAGAGVHHRRSSSTRSGGDLTLGLEPSEEARSPPLAPSEG 1080
 Qy 1081 AGSDVDPGDLGMAGAKGLOSLPTHDSPLORYSEDPVPLPSETDGVAPLTCSPPQPEYV 1140
 Db 1081 AGSDVDPGDLGMAGAKGLOSLPTHDSPLORYSEDPVPLPSETDGVAPLTCSPPQPEYV 1140
 Qy 1141 NOPDVRPQPSPPREGPLPARPAGATLERPKTSPGKNGVQVFAFGAVENBEYLTPQ 1200
 Db 1141 NOPDVRPQPSPPREGPLPARPAGATLERPKTSPGKNGVQVFAFGAVENBEYLTPQ 1200
 Qy 1201 GGAAPQHPPPAPSPAFDNLVYWDQDPERGAPSTFKGTPAENPEYGLDVPV 1255
 Db 1201 GGAAPQHPPPAPSPAFDNLVYWDQDPERGAPSTFKGTPAENPEYGLDVPV 1255

RESULT 2
 TVRNTU
 Protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
 C/Accession: A24562; A61204
 R/Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.
 Nature 319, 226-230, 1986
 A/Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
 A/Reference number: A24562; MUID:66118662; PMID:3945311
 A/Accession: A24562
 A/Molecule type: mRNA
 A/Residues: 1-1260 <BAR>
 A/Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746
 R/Menul, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen, C.
 Carcinogenesis 12, 1975-1978, 1991
 A/Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no 1-2-thiazolylformamide or N-methyl-N-nitrosourea.
 A/Reference number: A61204; MUID:92035293; PMID:1682063
 A/Accession: A61204
 A/Molecule type: DNA
 A/Residues: 637-663, 'V', 665-702 <MAS>
 A/Note: authors translated the codon GCA for residue 25 as Val
 C/Genetics:
 A/Gene: neu
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosph
 F/1-19/Domain: signal sequence #status predicted <SIG>
 F/20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

F:658-680/Domain: transmembrane #status predicted <TM>
 F:723-988/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:711,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:758/Active site: Lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 88.0%; Score 5997; DB 1; Length 1260;
 Best Local Similarity 87.9%; Pred. No. 2,5e-237;
 Matches 1105; Conservative 49; Mismatches 101; Indels 2; Gaps 2;

```

QY 1 MELAALCRWGLLALPPGASSTOVCTGTDMLRLPASPTHLDMLRLHYOGCQVQGNL 60
DB 4 MELAAMCRWGLLALPPGASSTOVCTGTDMLRLPASPTHLDMLRLHYOGCQVQGNL 63
QY 61 ELTYLPTNASLSTFODIOEVQGYVLIANNQVROPLORLIVRGTOLEFEDNYALAVLDNG 120
DB 64 ELTYVPANASLSTFODIOEVQGYVLIANNQVROPLORLIVRGTOLEFEDNYALAVLDNR 123
QY 121 DPLNNTTPVTGASPGGLRELQLRLSTLEILKGVLIQNPOLCYODITLMDIFHKNNOL 179
DB 124 DPDQNVAASTPGKRPBGLRELQLRLSTLEILKGVLIQNPOLCYODITLMDIFHKNNOL 183
QY 180 ALTLIDTNRBRACHPCSPMCKGSRGWSSSDCOSLRTVYAGGACACCKGPLPTDCCHQ 239
DB 184 APVDIDTNRBRACHPCSPMCKGSRGWSSSDCOSLRTVYAGGACACCKGPLPTDCCHQ 243
QY 240 CAAGCTPPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPBGRYTFGASCVTYAC 299
DB 244 CAAGCTPPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPBGRYTFGASCVTTC 303
QY 300 PNYVLTSDVSCITVCPFLNNOVTAEDGTORCEKCSKPCARVCYGLGMEHLREVRATSA 359
DB 304 PNYVLTSDVSCITVCPFLNNOVTAEDGTORCEKCSKPCARVCYGLGMEHLREVRATSD 363
QY 360 NIGEPACCKKIFGSLAPLPSPPDPAANTAPLOPEOLQVETLEETGYLYISAMPDSL 419
DB 364 NVGEPDCKKIFGSLAPLPSPPDPAANTAPLOPEOLQVETLEETGYLYISAMPDSL 423
QY 420 PDLSTVQNLQVITGRILHANGAYSITLQGLISWGLRLSRLBSGLALIHNTILCFVHT 479
DB 424 RDLSTVQNLQVITGRILHANGAYSITLQGLISWGLRLSRLBSGLALIHNTILCFVHT 483
QY 480 VPMADLFRNPHOALLHTANRPDE-CYBGLACQULCARHCHWPGPTQCVNCSQPLRG 538
DB 484 VPMADLFRNPHOALLHTANRPDE-CYBGLACQULCARHCHWPGPTQCVNCSQPLRG 543
QY 539 ECVBECRVLQGLPREYVNAHCLPCHECOPONGSVTCFPGPADQCAAHYDPPFCVA 598
DB 544 ECVBECRVLQGLPREYVNAHCLPCHECOPONGSVTCFPGPADQCAAHYDPPFCVA 603
QY 599 RCPBSPKPSLSTFODIOEVQGYVLIANNQVROPLORLIVRGTOLEFEDNYALAVLDNG 658
DB 604 RCPBSPKPSLSTFODIOEVQGYVLIANNQVROPLORLIVRGTOLEFEDNYALAVLDNR 663
QY 659 VGLLVVAVLGVVGIILKROQIKRTTWRRLLQETELVEPLTPSGAMPNOAMRIKET 718
DB 664 VGLLVVAVLGVVGIILKROQIKRTTWRRLLQETELVEPLTPSGAMPNOAMRIKET 723
QY 719 ELAKVAVLGSAGATGYKGIWIDGENVKIPVALIKVLRNTSPKANEILDEAYVAVAGV 778
DB 724 ELAKVAVLGSAGATGYKGIWIDGENVKIPVALIKVLRNTSPKANEILDEAYVAVAGV 783
QY 779 SPVYSLGLCLTSTVQVLTQMLPYGLLDHVENRGLSGODLLWCMQIAGKMSYLEB 838
DB 784 SPVYSLGLCLTSTVQVLTQMLPYGLLDHVENRGLSGODLLWCMQIAGKMSYLEB 843
QY 839 VRLVHRLDAARNVLYKSPNHYKTTDGLARLDIDETRYADGKVPYIKMALESILRR 898
DB 844 VRLVHRLDAARNVLYKSPNHYKTTDGLARLDIDETRYADGKVPYIKMALESILRR 903
QY 899 PTHQSDVMSGYVTVWELMTFGAIPYDGIAPAREIPDLBNGERLPQPPICITVDYMLMVK 958

```

```

DB 904 PTHQSDVMSGYVTVWELMTFGAIPYDGIAPAREIPDLBNGERLPQPPICITVDYMLMVK 963
QY 959 WVIDSECPREPRELVSERARRAPDORFVIONEDLGPASPLDSTFYSLIEDDMGDLY 1018
DB 964 WVIDSECPREPRELVSERARRAPDORFVIONEDLGPASPLDSTFYSLIEDDMGDLY 1023
QY 1019 DAEELYVPOQGFPCDDPAPGAGVNHHRNSSLTRSGGDLTLGLEPSEBEAPRSLAPS 1078
DB 1024 DAEELYVPOQGFPCDDPAPGAGVNHHRNSSLTRSGGDLTLGLEPSEBEAPRSLAPS 1083
QY 1079 EGASDVFDGDLGAGAKGLQSLPTHDSPLQORSEDPVTLBETDGVNAPLTSPOPE 1138
DB 1084 EGASDVFDGDLGAGAKGLQSLPTHDSPLQORSEDPVTLBETDGVNAPLTSPOPE 1143
QY 1139 YVNOQDVRPOPSPREBEPPLPAARPAATLERPKTSLPGKGVVDFVAFGAVENPEYLT 1198
DB 1144 YVNOQDVRPOPSPREBEPPLPAARPAATLERPKTSLPGKGVVDFVAFGAVENPEYLT 1203
QY 1199 PGGAAPOHPPPPAFSAFDNLTYWDDPPBRGAPPESTFKTPTAENPEYLGLDVPI 1255
DB 1204 PGGAAPOHPPPPAFSAFDNLTYWDDPPBRGAPPESTFKTPTAENPEYLGLDVPI 1260

```

RESULT 3
 I48161
 P-185. Precursor - golden hamster
 C/Species: Mesocricetus auratus (golden hamster)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
 C/Accession: I48161
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa
 Gene 140, 251-255, 1994
 A>Title: Cloning and activation of the Syrian hamster new proto-oncogene.
 A/Reference number: I48161; MUID:94193007; PMID:7908275
 A/Accession: I48161
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 11254 <RES>
 A/Cross-references: GB:016295; NID:9493236; PIDN:BA03801.1; PID:9747595
 C/Genetic:
 A/Gene: neu
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif

Query Match 87.9%; Score 5993.5; DB 2; Length 1254;
 Best Local Similarity 87.7%; Pred. No. 3.4e-237;
 Matches 1101; Conservative 57; Mismatches 96; Indels 1; Gaps 1;

```

QY 1 MELAALCRWGLLALPPGASSTOVCTGTDMLRLPASPTHLDMLRLHYOGCQVQGNL 60
DB 1 MELAAMCRWGLLALPPGASSTOVCTGTDMLRLPASPTHLDMLRLHYOGCQVQGNL 60
QY 61 ELTYLPTNASLSTFODIOEVQGYVLIANNQVROPLORLIVRGTOLEFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSTFODIOEVQGYVLIANNQVROPLORLIVRGTOLEFEDNYALAVLDNR 120
QY 121 DPLNNTTPVTGASPGGLRELQLRLSTLEILKGVLIQNPOLCYODITLMDIFHKNNOL 180
DB 121 DPLNNTTPVTGASPGGLRELQLRLSTLEILKGVLIQNPOLCYODITLMDIFHKNNOL 180
QY 181 LTLIDTNRBRACHPCSPMCKGSRGWSSSDCOSLRTVYAGGACACCKGPLPTDCCHQ 240
DB 181 LTLIDTNRBRACHPCSPMCKGSRGWSSSDCOSLRTVYAGGACACCKGPLPTDCCHQ 240
QY 241 AAGCTPPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPBGRYTFGASCVTYAC 300
DB 241 AAGCTPPKASDCLACHFNHSGICEIHCPLVYNTDTFESMNPBGRYTFGASCVTTC 300
QY 301 PNYVLTSDVSCITVCPFLNNOVTAEDGTORCEKCSKPCARVCYGLGMEHLREVRATSA 360
DB 301 PNYVLTSDVSCITVCPFLNNOVTAEDGTORCEKCSKPCARVCYGLGMEHLREVRATSA 360

```

```

Qy 361 IOEFAGCKKIFGSLAFLPESEFDGPASNTAPLOEQLOVFTLEITGYLYISAMPDLSL 420
Db 361 IOEFAGCKKIFGSLAFLPESEFDGPASNTAPLOEQLOVFTLEITGYLYISAMPDLSL 420
Qy 421 DLSVFNQLQVIRGILNNGAVSLTLQGLISWLGILSLRLGSGSLALIHNTNLCPVHTV 480
Db 421 DLSVFNQLQVIRGILNNGAVSLTLQGLISWLGILSLRLGSGSLALIHNTNLCPVHTV 480
Qy 481 PMOOLFNPQALHTNANREDECVBGLACHOCANRCHMGPRPTOCVNSOQLRQEC 540
Db 481 PMOOLFNPQALHTNANREDECVBGLACHOCANRCHMGPRPTOCVNSOQLRQEC 540
Qy 541 VEECRVQLGPREVYNARHCLPCHPECOPONGSYTCFEPADQCAAHYDPFCVARC 600
Db 541 VEECRVQLGPREVYNARHCLPCHPECOPONGSYTCFEPADQCAAHYDPFCVARC 600
Qy 601 PSQVPLSLYMPIMKYPDEEGMCQPCPINCSTHSCVDLDDKCCPABORASPLTISIIVG 660
Db 601 PSQVPLSLYMPIMKYPDEEGMCQPCPINCSTHSCVDLDDKCCPABORASPLTISIIVG 660
Qy 661 ILVVVTVGVVGLIIRKROCKIRKTYMRRLQETELVPLTPSGAMPNQAOMRLKETEL 720
Db 661 ILVVVTVGVVGLIIRKROCKIRKTYMRRLQETELVPLTPSGAMPNQAOMRLKETEL 720
Qy 721 RKVAVLSSGAGTGYKGIWIPDGENVKIPVAIKYLRNTSPKANKEILDEAYVYAGVSP 780
Db 721 RKVAVLSSGAGTGYKGIWIPDGENVKIPVAIKYLRNTSPKANKEILDEAYVYAGVSP 780
Qy 781 YVSHLLGICLTSTVQVLTQMLPYGCLLDHVENRGRUGSODLLMCMQIAKMSYLEDVR 840
Db 781 YVSHLLGICLTSTVQVLTQMLPYGCLLDHVENRGRUGSODLLMCMQIAKMSYLEDVR 840
Qy 841 LVHRDLAARNLVKSPNNVKTITDGLARLIDIDETRYHADGKVPKIMALESLIRRF 900
Db 841 LVHRDLAARNLVKSPNNVKTITDGLARLIDIDETRYHADGKVPKIMALESLIRRF 900
Qy 901 HOSDMSYGVTVVMEIWMETFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYMLVKCM 960
Db 901 HOSDMSYGVTVVMEIWMETFGAKPYDGIIPAREIPDLLEKGERLPOPICTIDVYMLVKCM 960
Qy 961 IDSECRPRFRELVESEFMAADPORFVVIQNEIDIGAPASPLDSTYRSLLEDMDGLVDA 1020
Db 961 IDSECRPRFRELVESEFMAADPORFVVIQNEIDIGAPASPLDSTYRSLLEDMDGLVDA 1020
Qy 1021 BEYLVPOQGFPCPPAPAGAGMHHRRSSSTRSGGGLUTGLBSESEKARSPLABSEG 1080
Db 1021 BEYLVPOQGFPCPPAPAGAGMHHRRSSSTRSGGGLUTGLBSESEKARSPLABSEG 1080
Qy 1081 AGSDVPFODDLGMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
Db 1081 AGSDVPFODDLGMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
Qy 1141 NQPEVRROPPLTPGCPPLPVPAAGATLERPKTLISGKGKGVVMDVTFPGGAVENPEYLVBR 1200
Db 1141 NQPEVRROPPLTPGCPPLPVPAAGATLERPKTLISGKGKGVVMDVTFPGGAVENPEYLVBR 1200
Qy 1201 GGAPFOPHPPPAPFAPDNLTYMDDPPERGAPSTFTGCTAENPEYLGIDVY 1255
Db 1201 GGAPFOPHPPPAPFAPDNLTYMDDPPERGAPSTFTGCTAENPEYLGIDVY 1255

```

RESULT 4

epidermal growth factor receptor precursor - human
 N:Contains: protein-tyrosine kinase (BC 2.7.1.112) erbB
 C:Species: Homo sapiens (man)
 C:Date: 15-Nov-1994 #sequence revision 27-Nov-1995 #text change 11-Jun-1999
 C:Accession: A00641, A25772, E30024, A30642, A33615, A23062, A05281, A60143, A33
 R:Ulrich, A.; Condeelis, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y
 rg, P.H.
 Nature 309, 418-425, 1984
 A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of

A:Reference number: A00641; MUID:84219729; PMID:6328312
 A:Accession: A00641
 A:Molecule type: mRNA
 A:Residues: 1-1210 <IDL>
 A:Cross-references: EMBL:X00588; NID:g31113; PIRN:CAA23240.1; PID:g757924
 A>Note: the authors translated the codon AAG for residue 540 as Asn
 R:Inili, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.
 Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
 A:Title: Characterization and sequence of the promoter region of the human epidermal gr
 A:Reference number: A25772; MUID:85270438; PMID:2991899
 A:Accession: A25772
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-29 <ISH>
 A:Cross-references: GB:M11234; NID:g181981; PIRN:AAA2370.1; PID:g553272
 R:Haley, J.; Whittle, N.; Bennett, P.; Kitchington, D.; Ulrich, A.; Waterfield, M.
 Oncogene Res. 1, 375-396, 1987
 A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification
 A:Reference number: S30024; MUID:88217333; PMID:3329716
 A:Accession: S30024
 A:Molecule type: DNA
 A:Residues: 1-29 <HA2>
 A:Cross-references: EMBL:X06370; NID:g31118; PIRN:CAA29668.1; PID:g31119
 R:Haley, J.D.; Waterfield, M.D.
 J. Biol. Chem. 266, 1746-1753, 1991
 A:Title: Contributory effects of de Novo transcription and premature transcript termina
 A:Reference number: A38672; MUID:91107677; PMID:1988448
 A:Accession: A38672
 A:Molecule type: DNA
 A:Residues: 1-29 <HA2>
 A:Cross-references: GB:A38425; NID:g181977; PIRN:AAA63171.1; PID:g553271
 A:Experimental source: carcinoma cell line A431-7
 R:Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; Me
 Nature 309, 806-810, 1984
 A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs
 A:Reference number: A00642; MUID:84245835; PMID:6330563
 A:Accession: A00642
 A:Molecule type: mRNA
 A:Residues: 'RCWERA', 150-187, 'KSVIAV', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-32
 ', 799-799, 'TD', 802-811, 'R', 813-942 <XUY>
 A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF recep
 R:Lin, C.R.; Chen, W.S.; Krutiger, W.; Stolarky, L.S.; Weber, W.; Evans, R.W.; Verna, I
 Science 224, 843-848, 1984
 A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplification
 A:Reference number: A43615; MUID:84196372; PMID:6326261
 A:Accession: A43615
 A:Molecule type: mRNA
 A:Residues: 713-964 <LIN>
 A:Experimental source: epidermoid carcinoma cell line A431
 R:Stimmen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.
 Biochem. Biophys. Res. Commun. 124, 125-132, 1984
 A:Reference number: A23062; MUID:85046483; PMID:6093780
 A:Accession: A23062
 A:Molecule type: mRNA
 A:Residues: 1028-1210 <STM>
 R:Weber, W.; Gull, G.N.; Speiser, J.
 Science 224, 294-297, 1984
 A:Reference number: A05281; MUID:84172183; PMID:6324343
 A:Accession: A05281
 A:Molecule type: protein
 A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>
 R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.
 J. Biol. Chem. 260, 5205-5208, 1985
 A:Title: Identification of residues in the nucleotide binding site of the epidermal gro
 A:Reference number: A60143; MUID:85182650; PMID:2985580
 A:Accession: A60143
 A:Molecule type: protein
 A:Residues: 740-744, 'X', 746-747 <RUS>
 R:McCzokowski, B.; Mosig, G.; Cohen, S.
 Nature 309, 270-273, 1984
 A:Title: ATP-stimulated interaction between epidermal growth factor receptor and superc
 A:Reference number: A38023; MUID:84191554; PMID:6325948
 A:Contents: annotation; receptor activity

A/Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
 R/Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.
 Cell 59, 33-43, 1989
 A/Title: Functional independence of the epidermal growth factor receptor from a domain
 A/Reference number: A3331; MUID:9000233; PMID:2790960
 A/Content: annotation, internalization signal
 C/Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor
 C/Genetics:
 A/Gene: GDB:EGFR
 A/Cross-references: GDB:120610; OMIM:131550
 A/Map position: 7p12.3-7p12.1
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
 F/1-24/Domain: signal sequence #status predicted <SIG>
 F/25-1210/Product: EGF receptor #status predicted <MAT>
 F/25-645/Domain: extracellular #status predicted <EXT>
 F/75-300/Domain: EGF receptor extracellular domain repeat <EET>
 F/390-600/Domain: EGF receptor extracellular domain repeat <EET>
 F/646-668/Domain: transmembrane #status predicted <TM>
 F/669-1210/Domain: intracellular #status predicted <INT>
 F/710-975/Domain: protein kinase homology <KIN>
 F/718-726/Region: protein kinase ATP-binding motif
 F/999-1046/Region: coiled-coil mediated internalization signal
 F/1047-1210/Region: inhibitory
 F/128,175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status predic
 F/745/Active site: Lys #status experimental

Query Match 46.5%; Score 3169; DB 1; Length 1210;
 Best Local Similarity 49.8%; Pred. No. 4e-122; Indels 116; Gaps 22;
 Matches 632; Conservative 177; Mismatches 345;

```

11 LLLALLPPGAA--STQVCTGDMKRLPASBETHLDMRLHYOCQVVOGNTLTYLPTN 68
14 LLAALCASALBEKKYCCQGSNTLTQLGFREDHFLSLQRMNNCEVLLGLLITTYQBN 73
69 ASLSPLDIOEVOGYVLIANQVQVPLQRLRYRGVQLPEDNYALAVLNDGPDLANTP 128
74 YDLSPLTIQVAVGYVLIANTVERIPLENLIIRGMYYENSYALAVLSYD----- 126
129 VTGASPGELRELQSLTEILKGGVLLQBNPOLCYODTILMKDIFHKNQALTLIDTNR 188
127 ---ANKTGLBELPRNLOEILHGVNRPNNPNCNVSIOQRDVSSDPLSNMSMDPQNH 183
189 SRACHPCSPACKSGRCWGESSEDCQSLTRTYCAGCA--RCKGPLPTDCHEQCAAGCTG 247
184 LSCCQKQDPCPCPNCSGAGBENQKTKIKIICAQCGCRCKGKSPDCCNQCAGCTGP 243
248 KHSDDCLALPHNSGICELCPALVTYNTDTFESMPNPEGRYTFGASCVTAAPNYLSTD 307
244 RESDCLVCRKFRDEATCKDCTCPMLLYNPPTYQMDVNPBGKYSFGATCVKKCPNRYVTD 303
308 VGSGTLVCLPHNQEVTAEDTGRCESKSPCARVCYGLGMEHLREVAATVTSANIQERAGC 367
304 HGSCVRAAGDSYEM--BEDGVRAKCKCBGRKVCNIGIGEFQDSLSINANTNKHFKNC 362
368 KCKFGSLAPLPESGDEDPASNTAPLOPQLOVPEFTLEITGYLTSAMPDLSPLSYFON 427
363 TSISGDLHILPVARGDSFTHTPPLDPOEIDLTKVKEITSTFLLIQAMPENRTDLAFEN 422
428 LOVIRGRILHNGAVSLTLQGLISWLGRLSLRELSGSLALIHNTTHCFVTVPMDFOLFR 487
423 LRIIRGTHQHGQSLAVSLNITSIGRLSKISDGVIIISGNKNCYANTIMKCLFG 482
488 NPHQALHTANRPDECVGSLAQHLCARHCHGPGFTQCVNCSQFLRQGECEYECRYL 547
483 TSGQKTKIINRGNKSCATQVCHALCSPECGPBRDVSGRNYSRGEBCVADCKKL 542
548 OGLPREYVNAHCLPCHPECOPOGNSVTCFGEBAQACACAAHYDDPPCVARCPSPYKPD 607
543 EGBREYVNSSECTQCCHPECLPQMANITCTGRGDNKICQAHYIDGHCVCYTCAGWNGE 602
608 LSTYPIWKFPDEBAGCPCPINCHSGVLDLQKCCPAQRAPLTSIISAVG---ILLV 664

```

```

DB 603 NNTL-VKRYADAGVCHLCHPNCCTYGTGPGLSGCPFNPKIP--STATGVALLLLV 659
QY 665 VLVGVVFGILIKRQOKIRKYTWRRLLQETELVEPLTPSGAMPQOQRILKETELARKY 724
DB 660 VALGIG--LPMRRRHVRRKTLRRLLQRELVESPLTPSGAPQOALLRIKETEFKIK 716
QY 725 VLSSGAFGYKGIWIPDGENVKI PVAIKYLRENTSFKANKELIDEAVVAGVSPVYSR 784
DB 717 VLSSGAFGYKGIWIPDGENVKI PVAIKYLRENTSFKANKELIDEAVVAGVSPVYSR 776
QY 785 LIGCLTSTVQLVTLQMPVGCCLDHYENRGRGSDOLMMQOIAQMSVLEEDVLR 844
DB 777 LIGCLTSTVQLVTLQMPVGCCLDHYENRGRGSDOLMMQOIAQMSVLEEDVLR 836
QY 845 DLAAKRVLSPNHVKITDGLARLLDIDETEVADGAKYPIKMMALLESILRRFTHQSD 904
DB 837 DLAAKRVLSPNHVKITDGLARLLDIDETEVADGAKYPIKMMALLESILRRFTHQSD 896
QY 905 VMSGYVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYIMVVKMIDSE 964
DB 897 VMSGYVWELMTFGSKPYDGI PASEISILKEGERLPQPICTIDVYIMVVKMIDAD 956
QY 965 CRPRRELVSEPRMARDPQRFVITO--NEDLPASPLDSTFYRSLDEDDMDGLVDAREY 1023
DB 957 SRPRFRLIEFSRMARDPQRFVITQDERMHLPSPTDSNFRALMDEDDMDVDADBY 1016
QY 1024 LVPQGFECFCDPAPAGAGMHHRRSSSTRSGGDLTLGLEPSEEAAPRSPLASEGAGS 1083
DB 1017 LVPQGFECFCDPAPAGAGMHHRRSSSTRSGGDLTLGLEPSEEAAPRSPLASEGAGS 1042
QY 1084 DVPDGLGMAAGLQSLPTHDSPLQRYSEDPVLPSET--DGYVAPLTCSPQPERYN 1141
DB 1043 N--NSTVACIDRNGLOSCPIKEDSFLQRYSSDPGALTEISIDDTFL-----PVPEYIN 1094
QY 1142 QPDVRRQPPSPRSGPLPAPACATL-----ERKTLSPGNQGVVQVAFAGAVENREY 1196
DB 1095 Q-----SVF--KRPASGVQNPVYHNOPLNAPSPSDPHYD--PHSTAGNREY 1138
QY 1197 L--TPQGAAPQPPAPFASPAFDLYWQD-----DP-----PERGAPSTPFG 1239
DB 1139 LNTVQ-----PFCVNSTFDSAPMAQKSHQSLDNDVYQDDPFRKAKRNGIRKG 1189
QY 1240 TPTAENPEYL 1249
DB 1190 S-TAENPEYL 1198

```

RESULT 5
 A53183
 A/Title: epidermal growth factor receptor precursor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 18-Jun-1999
 C/Accession: A53183; A43818; S24942; A28941; S43325; 149643
 R/Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Eard, H.S.; Jenkins, N.A.;
 Genes Dev. 8, 399-413, 1994
 A/Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor
 A/Reference number: A53183; MUID:94110986; PMID:8125255
 A/Accession: A53183
 A/Molecule type: mRNA
 A/Residues: 1-1210 <LUB>
 A/Cross-references: GB:U03425
 R/Avivi, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
 Oncogene 6, 673-676, 1991
 A/Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site
 A/Reference number: A43818; MUID:91232866; PMID:2030916
 A/Molecule type: mRNA
 A/Residues: 1-714 <AVI>
 A/Cross-references: GB:X59698
 R/Bislinger, D.P.; Serrero, G.
 submitted to the EMBL Data Library, June 1992
 A/Reference number: S24942
 A/Accession: S24942

A/Molecule type: mRNA
 A/Residues: 969-971, 'K', 973-1115, 'D' <EIS>
 A/Cross-references: EMBL:Z12608
 R/Heiser, G.C.; Gili, G.N.
 J. Biol. Chem. 263, 13152-13158, 1988
 A/Title: Epidermal growth factor receptor threonine and serine residues phosphorylated
 A/Reference number: A28941; MUID:8030814; PMID:3158233
 A/Accession: A28941
 A/Molecule type: protein
 A/Residues: 689-694, 'X', 696-704, 'L', 706-707, 989-992, 'XX', 995-996, 'X', 998-1000, 1002-1009, R/Hbb, M.L.; Dunn, A.R.; Alexander, W.S.
 submitted to the EMBL Data Library, April 1994
 A/Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor
 A/Reference number: S45325
 A/Accession: S45325
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-971, 'K', 973-1210 <VER>
 A/Cross-references: EMBL:X78987; NID:9488830; PIDN:CA55587.1; PID:9488831
 R/Perla, B.C.; Das, S.K.; Andrews, G.K.; Dey, S.K.
 Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
 A/Title: Expression of the epidermal growth factor receptor gene is regulated in mouse
 A/Reference number: I49643; MUID:93126380; PMID:7678348
 A/Accession: I49643
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 12-20, 22-132 <RES>
 A/Cross-references: GB:I06864; NID:9193001; PIDN:AAA53029.1; PID:9567201
 C/Genetics:
 A/Gene: EGFR
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keyword: ATP; growth factor receptor; kinase-related transforming protein; phosphoprotein
 F/1-24/Domain: signal sequence #status predicted <SIG>
 F/648-670/Domain: transmembrane #status predicted <TM>
 F/712-977/Domain: protein kinase homology <KIN>
 F/720-728/Region: protein kinase ATP-binding motif
 F/680, 695/Binding site: phosphate (Thr) (covalent) #status experimental
 F/697, 1070, 1071/Binding site: phosphate (Ser) (covalent) #status experimental
 F/993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
 F/1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
 F/1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 46.1%; Score 3145; DB 2; Length 1210;
 Best Local Similarity 49.7%; Pred. No. 3,86-121;
 Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;

11 LLLALLPFGA--STOYCTGTMDKRLPASPETHLMDRLHYOCGVQVQGLLEITLPTN 68
 14 LRLRLCAAGALBKKKVCQGTGNSRLTOLGTYEDHPLSLQRMYNCEVVLGSLLEITVORN 73
 69 ASLSFLDIOEVQGVYLIANQVROVPLRLIVRGTOLEPDNYALAVLDNGDPLANTTP 128
 74 YDSFLKTIQEVAVYLIANTVERIPLENQIIRGNALYENTYALAILSN----- 124
 129 VTGASPGGLAEQLRLSTLEIKGGVLIQRNPOLCYODITLWKDI---FKKNOLATLI 184
 125 -YGTNRGTLRELPRKNIQELIIGAVRFSNNPILCMQDITQWRDIQVONVFNMMQMDL-- 180
 185 DYNRSRACHPSPCKSGKRCWGSSESDQSLTRTYVCAAGCA-RCKGRLPTDCCHEQCAAG 243
 181 -QSPSSPCPKDPCSPNGSCWGGSENCQKLTIKIACQCSHRCKGNSPSCCHNQCAAG 239
 244 CTGPKHSDCLACLFNHSGLICELCPALVTYNTDTFESMPNREGRYTFGASCVTACPYNY 303
 240 CTGRRBDDCLVCGQFQDQATKDCPRLMLNPTTYQDVNPKGYSGACVKKCPKNY 299
 304 LSTVSGCTLVCPILNQEVTLEDTGQCKSKPCARVYGLGHEHLRFAVTSANIQ 363
 300 VVTHGSCVRAKCGDYEV-EDGIRCKKCDGCRKVCNIGIGEFQDTLSINATNIK 358
 364 FAGKKIKFGLSLAPLPSFSDQPSANTAPLOEQLOVFTLEIRGYLYISAMPDLPIS 423
 359 FKYCTALISGDLHILPVAKGDSPTKTPPLDRLRELILKTVKELTGFLTIQAMPDMDL 418

424 VFQNLQVIRGRIIANGAYSLTLOGLISWGLSLRELSGALLIHNTHLCFVHTVPM 483
 419 AFENLEIRGRTHQHQFSLAVVGLNITSLKSLKEISDGVLIISGNRLCYANTINWK 478
 484 QLRNPHQALHTANRPEDECVBGLAQHQLCARHGWGPGPQVNCQFRLROQCYEE 543
 479 KLFETPNQKTKIMNNAEKDCKAVNHCNPLCSSBGCGPPEPCVSCQVNSRGRECEK 538
 544 CEVLQGLPREYVNAHRLCPHPECQONGSVTCFGEPAADQVACAHYKDPFCVACPSG 603
 539 CHLLEGPFRFVENSECQHPECLQAMNITTCGRPDNICCAHITDGHCKYCTPAG 598
 604 VKEDLSYMPKPEDEBACQPCPINCTHSCVDLDKGCBAEORASPLTISAVGILL 663
 599 IMENNTL-VMKYADANNVCHLANCTYACAGPGLQGCCEWPSGPKIPISATGIVGLL 657
 664 VVVLGVVFGI-LKRRQKIKRTTMRLLQETELVEPLTSSGAMPNOAMILKTELK 722
 658 FIVV-VALGIGLPMRRRHIVRKETLRLLQERELVEPLTSSGEPNQAHLILKETEPRK 716
 723 VKVLGSAFGTYVKKGIWIPGENVKIPVALKVLRENTSPKANKELDEAVYMAVGSPYV 782
 717 IRLVGSARFTYVKKGLMIPGEKVKIPVALKELREASPKANKELDEAVYMAVDNHN 776
 783 SRLIGICTSTVOLVTOAMPYGLLDHVRNRRGLSSQDILNMCQIAGKMSYLEDVRLV 842
 777 CELIGICTSTVOLVTOAMPYGLLDHVRNRRGLSSQDILNMCQIAGKMSYLEDVRLV 836
 843 HEDLAARNTVKSPPNHYKITDPGLAARLDDERFVHADGKVPKXMALESILRRFHO 902
 837 HEDLAARNTVKSPPNHYKITDPGLAARLDDERFVHADGKVPKXMALESILRRFHO 896
 903 SDVWSYGVYTWELMTFCAKYDGIIPARIPDLLEKGRLLPOPICTIDVYIMYKMWID 962
 897 SDVWSYGVYTWELMTFCAKYDGIIPARIPDLLEKGRLLPOPICTIDVYIMYKMWID 956
 963 SECRPRELVSSESRMARDPQRFVITQ-NEDLGPASPLDSTFYSLLEDQMDGLVDAE 1021
 957 ADRPRKRELILFESQARDFQRYLVQGBRMHLPPDTSNPFYRALMDEEDMDVDAD 1016
 1022 EYLVPOGFCPPRPAFAGGVVHHRSSSTRSGGDLTGLSESEAPRSPLASGCA 1081
 1017 EYLVPOGFCPPRPAFAGGVVHHRSSSTRSGGDLTGLSESEAPRSPLASGCA 1042
 1082 GSDVFDGDLGMAKAGLSLPTDHPSPLOKRYSEDPVPLPSET---DGYVAPLTCSPQEPY 1139
 1043 TSN---NSTVYACINRNGSCRYKEDATLQRYSDPTGAVEDNIDDAFL-----PVREY 1092
 1140 VNQPDVRRPSPRSGPLPAPAPAGATL---BRPTLSFGKGVVVDVAFGAVENP 1194
 1093 VNQ-----SVP-KRPAQSVQNPVYHNPPLHAPVGRDLHYQN--PHSAVAVNP 1136
 1195 EYL-TPOGGAAPQHPHPAPSPAFDNLYWDQ-----DP-----PERGAPRPTF 1237
 1137 EYLVTAQ-----PTCLSSGFSNPAWLTQKSHQNSLDNPDYQDFFPKETKNGIF 1187
 1238 KGPRTAENPEYLGIDVP 1254
 1188 KG-PTAENAEYLVAVP 1203

RESULT 6
 T/CHLV
 epidermal growth factor receptor precursor - chicken
 N/Contains: protein-tyrosine kinase (EC 2.7.1.112) erdb
 C/Species: Gallus gallus (chicken)
 C/Date: 28-Feb-1986 #sequence revision 05-May-1995 #text change 04-Feb-2000
 C/Accession: A27720; A00643
 R/Lax, I.; Johnson, A.; How, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Vennet;
 Mol. Cell. Biol. 8, 1970-1978, 1988
 A/Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mo
 A/Reference number: A27720; MUID:88261272; PMID:3260329

A/Accession: A27720
 A/Molecule type: mRNA
 A/Residues: 1-1223 <LAX>
 A/Cross-references: GB:M20386
 R/Name: T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines, M.
 Cell 41, 719-726, 1995
 A/Title: c-erbB activation in ALV-induced erythroleukemia: novel RNA processing and pro
 A/Reference number: A00643; MUID:85228222; PMID:2988784
 A/Accession: A00643
 A/Molecule type: mRNA
 A/Residues: 585-1223 <NTL>
 A/Cross-references: GB:M10066
 A/Genetic: etbB
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor
 specific protein kinase
 F/1-30/Domain: signal sequence #status predicted <SIG>
 F/31-1223/Product: epidermal growth factor receptor #status predicted <MAT>
 F/31-654/Domain: extracellular #status predicted <EXT>
 F/81-307/Domain: EGF receptor extracellular domain repeat <EB1>
 F/397-610/Domain: EGF receptor extracellular domain repeat <EB2>
 F/655-677/Domain: transmembrane #status predicted <TM>
 F/678-1223/Domain: intracellular #status predicted <INT>
 F/719-884/Domain: protein kinase homology <KIN>
 F/727-735/Region: protein kinase ATP-binding motif
 F/136,202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #
 F/192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F/687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F/754/Active site: Lys #status predicted
 F/1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match 45.8%; Score 3122.5; DB 1; Length 1223;
 Best local similarity 48.7%; Pred. No. 3.1e-120;
 Matches 633; Conservative 175; Mismatches 345; Indels 145; Gaps 25;

QY 8 RWGILLALPRGAA-----STQVCTGTMKLRPASETHLMDLRHLVYGCQVVGNTL 61
 DB 13 RGAVALVLLGLVALCSAVEEKVCGGTNNLTQLGHEVDHFTSLQRYNNCEVTLSTLE 72
 QY 62 LTLPLTASLSFLDIDQVQVYLIANHQVQVPLQRLRYRGQVLPEDNALAVLNDGD 121
 DB 73 ITVYENHRDLFTLKTQEVAGVLIALMVDVIFLENIQIIRGNVLVDNSFALLVSTYH 132
 QY 122 PLNNTTPTVTSAPGSLRELQSLTEILKGVLIQRNPQLCYOPTILMKDIFHNKQNLAL 181
 DB 133 -NKTQ-----GLBELPKRLSEILNGVKSNNPKLCNMDVLANNDIIDSRL-PL 182
 QY 182 TLID-TYRSRACHPSCPMCKSRGSSBEDQSILRTVCAGGA-RCKGPLPTDCCHEQ 239
 DB 183 TVLDPAFNTLSGCPKCHPCTEDHCGAGGQNCQTLTKVTCAGQSGRGKVPDCCNQ 242
 QY 240 CAAGCTGPKASDCLACLFHNSGICELHCPALVTYNTDTPESMPNPGRYFGASCTTAC 299
 DB 243 CAAGCTGPRBSDDCLACRFRDADATCKDTPPLVYNPTTYQMDVPEKSYFGATCVAC 302
 QY 300 PYNVLTSDVSCITVCPLNQVTAADGTGRCCKSKCAVAVCYGLGMEHLREYRAVTS 359
 DB 303 PHNVVTVDHGCVRSCTDYEV-BENGVRCKCKDGLCSKVCNGIGELKGLISTNAT 361
 QY 360 NIOFACCKKIFGSLAFPSPDSPASNTAPLOPEQLQVETLEITGLYIYASMPSL 419
 DB 362 NIDSPKCKTNGVSVILPFAFLDAPTKTLPDPKLDVRYTVEISGFLILQAMPNA 421
 QY 420 PDLVSFQNLQVIRIRILHNGAYSLTLQGLGSWTGLSLRLSGSLAIHNTHLCVYHT 479
 DB 422 TDLVAFENLIRIRRTQKHGQYSLAVVNLKIQSLGSLKEISDGDILIMKNKMLCYADT 481
 QY 480 VPMQQLFRNPHQALLHTANREDGCGSLAQQLCANGHGMGPGPTQVACQFLRQGE 539
 DB 482 MNMNSLPRATQSKTKIIONKNDCTAORHVCDFLCSVVGCMGPGPFHCFSCRFPSRQKE 541
 QY 540 CVEBCRVLGSLPREYVNAHGLPCHPECPQNG--SVTCFGRPADQCACAHYKDPFPC 596

DB 542 CVKQCNILQGRPRFERBDSKCLPCHSBEVLQNSYAVNTTCSGPGPDHCKCKAHFIDGPHC 601
 QY 597 VAPRSGVPLDSTMPYWKPPDEBGACOPCINTCHSCVDDDDKCCPABQSRPLTSITS 656
 DB 602 VKACPAGVLGENDTL-VKTIADANAACOLCPNCTGCKGGLGECF--NSGKTPSTIA 657
 QY 657 AVV-GILLVVVLGVVFGILLIKRROOKIRKTYMRRLDQETELVEPLTPSGAMPNOQMRL 715
 DB 658 GVGGLLCLVVGIGIGYLRR-HYRKRLRLRLQRELVEPLTPSGEAPNOAHRLTL 716
 QY 716 KETELRKVTLGSGAFGVYVGIWIPDGENYKIPALIVLENSPKANKETLDEAYVA 775
 DB 717 KETELRKVTLGSGAFGVYVGIWIPDGENYKIPALIVLENSPKANKETLDEAYVA 776
 QY 776 GVGSPYRRLGICLTSTVOLVTOVLTQMPGCLLDHVRNKGRLSGODLNMCMQIKGMSY 835
 DB 777 SYDNPVCRLLGICLTSTVOLVTOVLTQMPGCLLDHVRNKGRLSGODLNMCMQIKGMSY 836
 QY 836 LEDVRLVHRDLAARVNLVKSBNHYKIDPGLARLLDIDETEVHADGKVPKIMALESTL 895
 DB 837 LEERLVRDLAARVNLVKTQHKITDFGLAKLGADEKEYHAGKVPKIMALESTL 896
 QY 896 RRRTHQSDWSYGVYTWELMTFGAKRPIDGIPAREIPDLKEXRLPPRITIDVYIM 955
 DB 897 HRITTHQSDWSYGVYTWELMTFGSKPYDGIABEISSEVLEKGERLPQPICTIDVYIM 956
 QY 956 VKCMNIDSECPRELVSESRMARDPQRFVVO-NEDLGPASLDSTFYSLLEDMD 1014
 DB 957 VKCMNIDSECPRELVSESRMARDPQRFVVO-NEDLGPASLDSTFYSLLEDMD 1016
 QY 1015 GDLVDAEYLVPQGFPCPDPAAGGVNHRHSSSTRSGGDLTLGLBSEEEAPRSP 1074
 DB 1017 EDIADAEYLVPQGF-----NSPST-----SRP 1042
 QY 1075 L-----APSGAGDVDPDGLGMAKGLQSLPTHDSPLORYSEDPVPLPSET--DCY 1127
 DB 1043 LLSLSATSNNSATNCID-----RNGQHPAREDSFYVORYSDPFGNPLESIDGF 1094
 QY 1128 VAPLTCRPOPEYVNPQRPQPSRPSRPLPAARPAGTLERPKTLSPGKGVVDVF-- 1185
 DB 1095 L-----PAREYVNO--LMPKPS-----TAAVQNDIYNNISLT 1125
 QY 1186 -----AFGAVENPEYLTPQGAAPQHPPEAPSPAFDNLVMDQ----- 1225
 DB 1126 ALSKLPMDSRQYNSHRAVDNPEYL-----NTNGSPFLAKTYVESSPYMIQSGNHQIN 1177
 QY 1226 -DPE-----RGAPSTFGKTPAENPEYGLDVP 1254
 DB 1178 LDNPYQODFLPNETKENGILLKVPALAEVYLRVAAP 1214

RESULT 7
 A47253 epidermal growth factor receptor, HER4 - human
 C/Species: Homo sapiens (man)
 C/Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
 C/Accession: A47253
 R/Plowman, G.D.; Culisecou, J.M.; Whitney, G.S.; Green, J.M.; Carlton, G.W.; Foy, L.; N
 Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993
 A/Title: Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal
 A/Reference number: A47253; MUID:93189574; PMID:8383326
 A/Accession: A47253
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: nucleic acid
 A/Residues: 1-1308 <PLO>
 A/Cross-references: GB:I07868; NID:9337359; PIDN:AB59446.1; PID:9337360
 A/Note: sequence extracted from NCBI backbone (NCBI:126842)
 C/Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; growth factor receptor
 F/716-961/Domain: protein kinase homology <KIN>
 F/724-732/Region: protein kinase ATP-binding motif

Query Match 44.1%; Score 3004.5; DB 2; Length 1308;
 Best Local Similarity 45.6%; Pred. No. 2,26-115;
 Matches 614; Conservative 183; Mismatches 375; Indels 175; Gaps 27;

9 MGLLLALPPGAA-----STVCTGTDMLRLPASPEHLDMLRLHYGSCOVVQGNLETTY 64
 8 WWWSLVLAAGTVPSPDSQVCACTENTLSLSLDEQYALRYTENCERVMGNLETTIS 67
 65 LPTNASISFLQDIOEVQSYVLIANOVROVPLQRLRIVRGTOLEFEDNYALAVLDNGPPLN 124
 68 IEHRDLSPLRSVRELVGYVALNQFPLYLENLRIRIGTKLVEDRALALIFLNYRNDG 127
 125 NTPPVTAASREGLELQRLSLTEILKSGVLIQRNPOLCYODTIIMKDI FHKNNQALTLI 184
 128 NF-----GLOELGKLTLEILNGGVYDQNKFLCYADTIIMODIVRNPMPSNLTLY 178
 185 DTRSRACHPCSPKCSKRGWSSSDQSLTRVYVCAAGC-ARCKGPLTDCCHQCAAG 243
 179 STNGSSCGGRHKSCTG-RCWGPTEHNCQYLITRYVCAQCQGRCYGVPVSDCCHRECAAG 237
 244 CTGPKHSDCLACTAFHNSGICELHCPALVTYNTDTPESSMPNREGRYTFGASCCTACPYNY 303
 238 CSGKPDIDCRACNNPNDSGACVTCQPTFVYNPTTFQLEHNPNAKITYTGACVCKCPHNF 297
 304 LSTDVSGCTLVCPHNOBVTAEDETQRCCKSKPCARVYGLGMEHLREVAVTSANIQE 363
 298 V-VDSSSCVACSPSSKMEV--EENGIMKCKPCTDICKPKACDGICTGSLMSAQTVSSNDK 355
 364 PACGKTI FGSIALPLPESPGDPAANTAPLEQOQVPTLEITGYLYISAMPDSDLS 423
 356 FINTKTNGLIFLVGTIGHDPYNAIEAIDCKLVNFTVEITGFLMIQWSPMPMDPS 415
 424 VFQNLQVIRGRILHNGAVSLTLQGLISWLGSLRSLEKSGSLALIHNTHLCPYATVPMD 483
 416 VFSVLVITGKRVLYSGSLTLIKQGGITSQFGSLKEISAENITYITNSNLCYHNTMT 475
 484 QLPNPQALHTANRPEDECVGSLACHQICANHCCKGPPPTQVNCQSLRGQSCYEE 543
 476 TLFSTINQRIYIRNRKAKENCTAGAMVCNHLCSGSGCPGPDCLSCRRFSRGICIES 535
 544 CRVLQGLPREVYVNRHCLPCHREPCO-ONGSVTCGPRADQVCAHAKDPPFCARPS 602
 536 CNLTDGEREPENSGICVECDPCKMEBDGLTCHGEPDPDCTKSHKQDGNCEKCPD 595
 603 GVXKDLSTYPIWKPEDEGACOPRPINCTHSCVLDLDKGC-----PABORASPL 651
 596 GLQGANF--IFKXADPRBCHPCHPNTQGCNPTSHDCIYYPWTGHTSLPQAR----- 649
 652 TSITSAVV--GILLVVVLGVVFGILIKRQOKIRKTYNRLLQETELVEPLTPSGAMPQ 709
 650 TPLIAGVIGGLFLIVIVGLTFVAVYVRKSIK-KKRALRRFL-ETLEVEPLTPSGTAPNQ 707
 710 AOMLILKTELRKTVYVGGAGFAGTVYKGIWI PDGNNVYI PAIVIKLRNTSPKXNKELD 769
 708 AQLILKTELRKRVYVGGAGFAGTVYKGIWIPBETVYI PAIVIKLINTTGKAVNEPD 767
 770 EAYVAGVSPYVRLIGLICLTSTVOLVTQMLPYGCLLDHYRENGRGLSODLLNMCQI 829
 768 EALIMAGNDHHLVRLIGVCSPIQLVLTQMLPHGCLLEVENHENDNGSOLLNMCQI 827
 830 AKGMSYLEDVRLVHRDLAARVLYKSPNHVKITDPGLARLLDIDETEHADGKVPYKIM 889
 828 AKGMSYLEERLVHRDLAARVLYKSPNHVKITDPGLARLLDIDETEHADGKVPYKIM 887
 890 ALBSILRRFTHOSDWSYGVTVLMEITFGAKPYDGI PABRI PDLLEKGBRLPORPCTI 949
 888 ALBSILRRFTHOSDWSYGVTVLMEITFGAKPYDGI PABRI PDLLEKGBRLPORPCTI 947
 950 DVYVIMYKCMNIDSECRPRELVESEFARMADPORFVYI QNED-LGSPPLDSTFYRSL 1008
 948 DVYVIMYKCMNIDSECRPRELVESEFARMADPORFVYI QNED-LGSPPLDSTFYRSL 1007
 1009 LEDDMGDLVAEYLVLPQGRPCPDPAFGAGVYVHNRSSSTRSGGGLTLGLEPSEE 1068

Query Match 39.7%; Score 2708; DB 1; Length 1166;
 Best Local Similarity 45.5%; Pred. No. 2,36-103;
 Matches 576; Conservative 164; Mismatches 389; Indels 138; Gaps 28;

1008 LDEEDLEDMWDAEYLVLP-QAFNIPP-----LYTSRARIDSRS-----EIGSPPPA 1055
 1069 EAPRS-----PLAP-SEGAGSDVFPDGLGMGAAGLQ 1099
 1056 YTPMSGNQPYVRRDGAFAEGGVSVYPRAPSTITIEAIVAGATRAEIPDSCNGTLRKPV 1115
 1100 SLPTHDSPLQRYSEDPVPLPS-----ETDGYVAPLTCSPQPEYVNOQDVPRPQPPSP 1152
 1116 APHQEDSSYQYSADPTVAPERSPRGELDEBSGYMTPMRKPKQOEYLNLYVE----- 1167
 1153 REGPLPAPRAGATLERPKTLSPKNGVYDVAFGCAVENEPEYLTPOCGAARPOHPPPA 1212
 1168 -ENFVSRR-----KNQDLO-----ALNDPEYHNASNG-----PPKA 1198
 1213 -----FSPAFDLYVWDDPPERGA--PPST 1236
 1199 EDEYNEPPLYANTPANTLGAKEYIAKNNILSMPEKAKAFDNDPYNHSLPFRSTLQHDPY 1258
 1237 FKGTPT-----AENPEYL 1249
 1259 LQYSTKCYFKQNGRIRPIVAENPEYL 1285

RESULT 8
 506142
 protein-tyrosine kinase (BC 2,7,1,112) mrk-y precursor - southern platyfish
 N:Alternate names: epidermal growth factor receptor homolog; kinase-related transformin-
 C:Species: Xiphophorus maculatus (southern platyfish)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Feb-2000
 C:Accession: S06142; S12809
 R:Witbtdrod, J.; Adam, D.; Maltischeck, B.; Maenele, W.; Raulf, F.; Telling, A.; Robert
 Nature 341, 415-421, 1989
 A:Title: Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu 10
 A:Reference number: S06142; MUID:90015140; PMID:2797166
 A:Accession: S06142
 A:Molecule type: DNA
 A:Residues: 1-1166 <MIT>
 A:Cross-references: EMBL:X16891; NID:965290; PIDN:CAA34770.1; PID:965291
 R:Adam, D.; Maenele, W.; Scharf, M.
 Oncogene 6, 73-80, 1991
 A:Title: Transcriptional activation of the melanoma inducing Xmrk oncogene in Xiphophor
 A:Reference number: S13807; MUID:91125882; PMID:1846957
 A:Accession: S13809
 A:Status: preliminary; translation not shown
 A:Residues: 821-1025 'N', 1027-1098 'A', 1100-1166 <ADA>
 A:Molecule type: DNA
 A:Cross-references: EMBL:X56319; NID:965284; PIDN:CAA39763.1; PID:965285
 C:Genetics:
 A:Gene: mrk
 A:Map position: Y
 A:Intons: 872/3; 898/1; 947/1; 979/3; 1025/3; 1056/1
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor; phosphotransferase; transmembrane protein; tyr
 F:1-53/Domain: signal sequence #status predicted <SIG>
 F:126-1166/Product: kinase-related transforming protein (Tn) #status predicted <MAT>
 F:707-972/Domain: protein kinase homology <KIN>
 F:715-723/Region: protein kinase ATP-binding motif

Query Match 39.7%; Score 2708; DB 1; Length 1166;
 Best Local Similarity 45.5%; Pred. No. 2,36-103;
 Matches 576; Conservative 164; Mismatches 389; Indels 138; Gaps 28;

4 AALCRWGLLALPPGAAT-----QVCTGTDMLRLPASPEHLDMLRLHYGSCOVVQGN 59
 8 AALQO--LLVLSLRCCSTDPDRKVCQGTNSQNTM---LDNHLYKMKKMSGCVNVLLEN 62
 60 LEITLPTNASLSFLQDIOEVQSYVLIANOVROVPLQRLRIVRGTOLEFEDNYALAVLDN 119
 63 LEITYQENDLSFLQSIQEVGYVLIAMNEVSTIPLVNLRLRIGONLYEGNFTLLVMSN 122
 120 GDPINNTTPVTAASPGGLRELQRLSLTEILKSGVLIQRNPOLCYODTIIMKDI FHKNNQ 179

```

Db 123 YOK-NPSSP--DVYQVGLKQLOLSNLTLLSGGVKSHNPCLCNVETINMMDIIDXTSNP 179
Qy 180 ALTLIDTNRSPACHPSPMCKSGSRGWESSBDCQSLTRTVACAGGC--ARCKPLPTDCCH 238
Db 180 TNNLIPHAEROCQKCDHCVCANSCWAPRGHCQKFTKLLCAEQCRRCRGRPRIDCCNE 239
Qy 229 OCAAGCTGKXSDCLACHFNHSGICELCPALVTYNTDTFESMPNBSRYTFGASCVA 298
Db 240 HCAGCTGPRATDCLACRDFNDGTCKDTCPPKIDYVSHQVADPNKIKYTFGAACVKE 299
Qy 229 CRYNYLSTDVSGCTVLCPLHNOEVTAEDETPORCEKSKCARCYVGLMEHLREVAVVS 358
Db 300 CFSNTVYVE--GACVNSCSAGMLEVD--ENGRSKSKPCDGVCPKVCDDIGISLNTIAVNS 357
Qy 359 ANIOEPAGCKKIFGSLAFPLPESFGDPASNTAPLOEOLQVETLEITGYLYISAMPSS 418
Db 356 TMRBFSNCTKINGDILNRNBSGDPHKIGTMDPEHMLNLTVEITGYLYIVMMPEN 417
Qy 419 LPDLISVFNQVIRGRILHNGAYS--LTQGLGISWGLRLSRLBSGLALIHNTLCEV 477
Db 418 MTSLSVFNQVIRGRILHNGAYS--LTQGLGISWGLRLSRLBSGLALIHNTLCEV 477
Qy 478 HTPVPDQPLRNHQAHLHNRNBSGDPHKIGTMDPEHMLNLTVEITGYLYIVMMPEN 537
Db 478 HTPVPDQPLRNHQAHLHNRNBSGDPHKIGTMDPEHMLNLTVEITGYLYIVMMPEN 537
Qy 538 QECVSECRVLOGLPREYVNAHRLCPHPCQOPONGSVTCFPGPADQCVCAHYKDDPPCY 597
Db 530 GRCVASCNLQGBRPAQVDRGCVQCHQCLVQDTSITCTGPRPANCSSAHHQDDPQCI 589
Qy 558 ARCPGKVPDLISYMPIWKPDPBEGACQPCPINCTHSCVDLDKXCPAEGPASPISIIA 657
Db 590 PRCPFGIILDDGDTL--IMKYADKMGQCPQCHQNTQCCSPGLSGCGD--IVSHSLAVGL 647
Qy 658 VGLILLVAVLVGVLIRBROOKIKRYMRLLOTELYEPLTPSGAMPNQAQMLKE 717
Db 648 VGLILLVAVLVGVLIRBROOKIKRYMRLLOTELYEPLTPSGAMPNQAQMLKE 717
Qy 718 TELRKVYVLSGAGFTVYKGIWIPDGENYKIPVAIKVLRNTPSKANKILDEAYVAGV 777
Db 707 TEPKQDRLVSGAGFTVYKGIWIPDGENYKIPVAIKVLRNTPSKANKILDEAYVAGV 766
Qy 778 GSPYVSRLLIGLCTSTQVLTQMLPGCLLDHVRNRRGLSGDILLNMCQIAKNSYLE 837
Db 767 DBHPCRLILGICLTSVAVQVLTQMLPGCLLDHVRNRRGLSGDILLNMCQIAKNSYLE 826
Qy 838 DVLVYHRLDLAARVLYKSPHYKITTDFGLARLLDIDETRYHADGGVPIKMALESILR 897
Db 827 ERLVYHRLDLAARVLYKSPHYKITTDFGLARLLDIDETRYHADGGVPIKMALESILR 886
Qy 898 RHTHSDVWSYGVWELMTFGAKPYDGI--PAREIPDLLEKGERLPOPICTIDVYVIMYK 957
Db 887 TYTHSDVWSYGVWELMTFGAKPYDGI--PAREIPDLLEKGERLPOPICTIDVYVIMYK 946
Qy 958 CMMDISECRPRRELVSERSMARADPQRFVIONEDLGPASPLDSTFYRSLLEDDMGDL 1017
Db 947 CMMDISECRPRRELVSERSMARADPQRFVIONEDLGPASPLDSTFYRSLLEDDMGDL 1001
Qy 1018 VDAEYLVQOQGFPCPDPAAGAGVHHRSSSTSGGDLTLGLEPSEBBAAPSLAP 1077
Db 1002 VDAEYLVQOQGFPCPDPAAGAGVHHRSSSTSGGDLTLGLEPSEBBAAPSLAP 1077
Qy 1078 SRGASDVVDGDLGMAKGLQSLPTHDPSPLORYBEDPTV--PLPSETDGYVAPLCSQ 1136
Db 1025 PFGH-----PVRENSITLRNLSIDPQNALKEKLDGH----- 1055
Qy 1137 PEYVQOPDVRPOP-----PSPRE-----GFLP--AARPAQATLBRPKTSLSPKNGVVD 1183
Db 1056 -EYVQOPSETSRSLSDINPNVEDLIDMGVPSVSSQBAETNFSRPEYVANTQNL----- 1111
Qy 1184 VPAFGAVENPEYLTQGAAPQHPHPAPAPFNDLTYWDDPPEBGAPESTFKCTPTA 1243
Db 1112 PLVSGSGMDPPY--QAG-----YQAAF-----LPQTGALTGNGMFLPAA 1149

```

```

Qy 1244 ENPEYLG 1250
Db 1150 ENPEYLG 1156

RESULT 9
A:Accession: A36223
A:Species: Homo sapiens (man)
A:Date: 04-Oct-1991 #sequence_revision 13-Jan-1993 #text_change 17-Nov-2000
A:Accession: A36223; 159164
A:Title: Isolation and characterization of ERBB3, a third member of the ERBB/epidermal
A:Reference number: A36223; MUID:90083234; PMID:2687875
A:Accession: A36223
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1342 <KRA>
A:Cross-references: GB:M29366
R:Plowman, G.D.; Whitney, G.S.; Neubauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990
A:Title: Molecular cloning and expression of another epidermal growth factor receptor-
A:Reference number: 159164; MUID:90311312; PMID:2164210
A:Accession: 159164
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-559, 'G', 561-957, 'F', 959-1063, 'G', 1065-1342 <RSS>
A:Cross-references: GB:M34309; NID:9183990; PIDN:AAA35979.1; PID:9306841
C:Gene: ERBB3; HER3
A:Gene: GDB:ERBB3; HER3
A:Cross-references: GDB:119880; OMIM:190151
A:Map position: 12q13-12q13
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C:Keywords: ATP; phosphotransferase
F:707-972/Domain: protein kinase homology <KIN>
F:715-723/Region: protein kinase ATP-binding motif

Query Match 35.7%; Score 2432.5; DB 2; Length 1342;
Best Local Similarity 40.7%; Pred. No. 4,5e-92;
Matches 533; Conservative 191; Mismatches 458; Indels 129; Gaps 32;

Qy 10 GILLALPPGAA--STQVCTGDMKRLPASPEHMLNLYQCGVVOGNLELYLPT 67
Db 11 GILFSLARSEVNGAGVCPGTLNGSLVGDENQVTLKYKEREVEVWNGLEIYVLGH 70
Qy 68 NLSLFLQDIOGVGYVLIANQVQVPLQRLIRVGTQLFEDVVALAVLDNGDPLNNTT 127
Db 71 NADLSFLQWIRVYTGVLVANNBFTLPLNLRVVGTVGQVYDGFAPV-----LNVT 125
Qy 128 PTTGASPGRLRLQLRSLTEILKGVLIQNRNQLCYODITLMWDIHHKNNQLALTLDTN 187
Db 126 ---NSSHALRQLRLQLTEILSGVYIEKNDLCMDITDMVDIVRDR--AEIVKD 178
Qy 188 RARAPCPSPMCKSGSRGWESSBDCQSLTRTVACAGGC--ARCKPLPTDCCHOCACCTG 246
Db 179 NGRSCPCHVEVCKG--RCPGSEDCQTLTKTICAPQNGHCFEPNPOCHDCACGCSG 237
Qy 247 PHSDDLACHFNHSGICELCPALVTYNTDTFESMPNBSRYTFGASCYVAPYNYLST 306
Db 236 PDLTDFACRHNFDNSACVPRGPQLVYKGLFQLBPNPHTKQYGVGVCAACPFHNV--V 296
Qy 307 DYGSCVLPRLNQBVTADGTQRCCKSKPCARVYCYGLGMHLREVRVATSNIOEFAG 366
Db 297 DQTSVCRACPDPKMEVD--KNGKMKCBPCCGLCPKACBEGSG--SRFQTDSSNIDGFVN 353
Qy 367 CKKIFGSLAFPLPESFGDPASNTAPLOEOLQVETLEITGYLYISAMPDLSVQ 426
Db 354 CTKLGNLDFLTGLNGDPWHKIPALDPEKLVFRVRELTGLNTQSNPPHNPVSVFS 413
Qy 427 NQVIRGRILHNGAYS--LTQGLGISWGLRLSRLBSGLALIHNTLCEVATVPWDOL 485

```

```

Db      414 NLTTIGRSILYNRGFSLLIMKNLNTSLGFRSLKEISAGRIYISANRQLCTHHSIMNTKV 473
Qy      486 FRNPQALLHTA-NRPEDECVGEGLAGLCHOLCARGHCWGPPTQCVNCSQPLRQCEVEEC 544
Db      474 LRQPTRELRDILKNNRPRRDCVAEGKVCDDLCSGSGCGPFGQCLSCRYN SRGGVCYTHC 533
Qy      545 RVLGGLPREVYVNAHCLPCHPECOPONGSVTCGPRADQVACAHYDPPFCVACRSGV 604
Db      534 NPLNGSPREPRHMAECFSCPECPMGATCNGSGDTCQAQCHFRDGHCVSSCPHGV 593
Qy      605 KPLDSVMPKIPDEBEGACOPCPINCHSGVDLDDKCCPRAQRA---SPLTSLISAVVG 660
Db      594 LG--AGGIYKYVDVQNECRPCHECTQCGKPELDQCLQGLVLTGKTHLTALTYIAG 651
Qy      661 ILLVVLGVVFGILIKRQOKIR-KYTMRLLOTELVEBPLTPSGAMPNOAKRILKETE 719
Db      652 --LVVIFMVLGGTFVLVYRGRIQNKRAMRYLERGSEIPLDPS-BKANKVLAIRFETE 708
Qy      720 LRKVKTLGSGARCTVYKGIWIPDGENYKIPVALKYLBENTSPKANKELDEAYVMAGVS 779
Db      709 LRRLKVLGSGVFTVHKGVMIPEGESIKIPVCIKVIDKSGRQSFQAVTDMLAGSLDH 768
Qy      780 PYVSRLLIGLCTVTVOLVTQMPYGLLDHYRNRGRGLSGODLLNMCQIAKGSYLEDV 839
Db      769 AMLVRLIGLCPGSSLDLVLYPLGSLLDHYRQRRGALGPOLLNMGVQLAKGTYLEEH 828
Qy      840 RLVRDLAARNVLVKSFNHYKITDFGLARLLDIDETRYHADGGKVPKIMNALSILRRF 899
Db      829 GNVHRLAARNVLLKSPSQVQVADPFGVADLLPPDDQLLVSEAKTPIKMALLSIHKGX 888
Qy      900 THOSDVSVYVYTWELMTFGAKPYDGIIPABEIPDLKESGRLPOPPCTIDVYIMVKW 959
Db      889 THOSDVSVYVYTWELMTFGAEPYAGRLAEVPLLEKGERLAQPCITIDVYVWVWKW 948
Qy      960 MIDSECRPRELVSEPRARDPORFVVIQNEIDLGA---SPLDSFFYSLLLEDDMGD 1016
Db      949 MIDENRPTKELANETTRARADPRVLTIKRS-GEGLAPGEPHGLTKKLEKVELP 1007
Qy      1017 LVDAEYLVPQGFPCDPAAGAGVNHRRSSSTRSGGDLTLGLEP-SEEBAPRSP 1075
Db      1008 ELRLDLDAEED-----NLATTLGSLSLPYGLTINRPRGSL 1048
Qy      1076 APSEGAGSDVFDGDLGMAKGLQSLFTHD-PSFLQKYSDDPYVPLP-----SETDGYV 1128
Db      1049 SPSSGY-MPNQGNLGBSCQESAVSGSERCPREVSLH-----PMERGCLASESSEGHV 1101
Qy      1129 A-----PLTCSPOPE---YVNPQDVPRPSPRSGP-----L 1157
Db      1102 TGSEBALQEVSMCRSRSRSPRPRDSDAHSQSHLTPVTPLSFPGLEBEDVNGYVM 1161
Qy      1158 PAARPAATLERPRTLSR-GKNGVY-----KQVAFAGAVENPEVLTPOGGAAPQHP 1210
Db      1162 PDTHLGTGPRSGRTSLSSVGLSVLTGEBEED-----BEYEVNRRRRRSP-PHP 1212
Qy      1211 PARSAPADNLYND-----QDPRKAPSTFKGTPLAENPEYL 1249
Db      1213 RPSLEBELGYEYDVGSDLSASLGSTQSCPLHPVPIPTAGTTPDEDEYEM 1263

```

RESULT 10

```

Jc4387
epidermal growth factor receptor homolog precursor - rat
N/Alternate names: ErbB3 protein; HER3 protein
C/Species: Rattus norvegicus (Norway rat)
C/Date: 17-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 13-Nov-1998
R/hellier, N.J.; Kim, H.H.; Greaves, C.H.; Sierke, S.L.; Koland, J.G.
Gene 165, 279-284, 1995
A/Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protein.
A/Accession number: Jc4387; MUID:96096535; PMID:8522190
A/Molecule type: mRNA
A/Residues: 1-1339 <HEL>

```

```

A/Cross-references: GB:U29339; NID:9915389; PID:9915390
A/Experimental source: Liver
A/Note: The authors translated the codon AAC for residue 369 as Thr and GTT for residue
C/Comment: This protein is a functional heregulin receptor that transduces signals to ti
C/Genetic:
A/Genes: ErbB3
C/Superfamily: Unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-1339/Product: epidermal growth factor homolog #status predicted <MAT>
F/640-659/Domain: transmembrane #status predicted <TMW>
F/705-970/Domain: protein kinase homology <KIN>
F/713-721/Region: protein kinase ATP-binding motif
F/939,1051,1156,1196,1219,1257,1259,1273,1286,1325/binding site: phosphate (Tyr) (
Query Match 34.4%; Score 2347.5; DB 2; Length 1339;
Best local Similarity 40.8%; Pred. No. 1,3e-88;
Matches 524; Conservative 170; Mismatches 434; Indels 155; Gaps 34;
Qy      3 LAAICRWMGLLALLPRGAA--STQVCTGDMKRLRPASBETHLDMRLHYQCGQVQGN 59
Db      7 LQVLC---FLSLARGSEWNSQAVCPGTLNGLSTGADNQGQVLYKLYKECEVMGN 62
Qy      60 LELTYPTMASLFLDIOGVQVYLIANNQVQVPLQRLRYRGTQOLFEDNYALAYLDN 119
Db      63 LELVTHGNADSLFLQWIREVAYVLANNEFSVPLPMLRYRGTQVYDGKFAIYV-- 120
Qy      120 GDLPLNTPPTVPGASPGGLRELQRLSTELIKGVLIOBNPOLCYQDTILMKDIFHKNNQL 179
Db      121 --LANTY-----NSIALNQLKFTQLTLEISGVYIEKNDKLCMDITIDRDIYVR-- 170
Qy      180 ALTLIDTNSRACHPCSPCKGSRGWSGSESDQSLTRIVCAGG-ARCKGLPTDCHE 238
Db      171 GAETIVKNNGANPCPGHEVCKG-RCWGPGEDDCQILTKTICAPQCNCRCPGPNQCCHD 229
Qy      239 QGAAGCTGPRHSCLACLFHNSGICELHCPALVYNTDPEESMPNDEGVTTRASCVTA 298
Db      230 EGAGSGSGPDQDTCFACRRFNDSGACVPRCPRELVNKLTFQLEPNHTTYOYGGVAVAS 289
Qy      299 CFYNYLSTVDSCTVLCPLHNOETVADGTQRCCKSKPCARVCYGLMEHLREVRVATS 358
Db      290 CFHNPV-VQOTFCVRAACPRPKAEVD-KHGLKMEPCGGLCPKACEGSG--SRVQVDS 345
Qy      359 ANIQEFAAGCKIFGSLAFPESEFDGPASNTABLOPEQLOVFETLEITGYLYSAMPDS 418
Db      346 SNIDGVNCTKILGNDFLTGLAVDPBMHKIPALDEKLANVFVREITGYLNIOSWPH 405
Qy      419 LPLSLFQNLQVIRKGLINAGYS-LTQGLGISWGLRSLRELGSGLALINHTHLCFV 477
Db      406 MNPVSFNSLTTTGSLYNRGFSLLIMKNLNTSLGFRSLKEISAGRYISANQQLCYH 465
Qy      478 HTVPMQLFRNPQALLHTA-NRPEDECVGEGLAGLCHOLCARGHCWGPPTQCVNCSQPLR 536
Db      466 HSLMTRLRKGSEERLDIKYDRPLGELABGKVCDBPLCSGSGCGWAPQCLSCRYNKR 525
Qy      537 GQCEVEECRYLGLPREYVNAHCLPCHPECOPONGSVTCGPRADQVACAHYKDRPFC 596
Db      526 EGVCTVHNCNPLQGEPRFVHAQCFSCHEPCLPMBGISTYNGSGSDACACAHFRDPSHC 585
Qy      597 VARCPSGVKPLSYMTMKPRDEBEGACOPCPINCHTSC--VDLDDKCPRAEQRAPSLTI 654
Db      586 VNSCPHGILG-AKGPYKYPDAQNBRCPCHECTQCGKPELDQCLQGLAEVLSMKPHLV 643
Qy      655 ISAVGILLVVLGVVFGILIKRQOKIR-KYTMRLLOTELVEBPLTPSGAMPNOAKMR 713
Db      644 IAYTVG--LAVILMILGSLYKRGRIQNKRAMRYLERGSEIPLDPS-BKANKVLAIR 700
Qy      714 ILKTELKRVKVLGSAFGTVYKGIWIPDGENYKIPVALKYLBENTSPKANKELDEAYV 773
Db      701 IFKTELKRLKVLGSGVFTVHKGVMIPEGESIKIPVCIKVIDKSGRQSFQAVTDMHLA 760
Qy      774 MGVGSPYVSRLLIGLCTVTVOLVTQMPYGLLDHYRNRGRGLSGODLLNMCQIAKGM 833

```

Db 761 VGSLSAHVRLGLCGSSLDQVTOYLPGLSLDHVKHRETLGQOLLINMGVOIAKGM 820
 Qy 834 SYLEDVRLVHRDLAANVVKSPNHVKITDFGLARLIDIDETEHADGKVPIMKMALES 893
 Db 821 YLIEBHSVHRDLALNNVMSKSPQVADLPPDDKQLHSEAKTPIKMMALLES 880
 Qy 894 ILRRRTTHOSDWSYGVTTWELMTFGAKPYDGI PAEIPDLLEKGERLPORPITCTIDVYM 953
 Db 881 IHFGKTHOSDWSYGVTTWELMTFGAEPYAGLRLMEIPDLLEKGERLAPQICITIDVYM 940
 Qy 954 IAWKCMIDSECPREELVSEFSRMAARDORFVUIONEDLGASPLDSFTFYSILIEDD 1013
 Db 941 IAWKCMIDSECPREELVSEFSRMAARDORFVUIONEDLGASPLDSFTFYSILIEDD 997
 Qy 1014 MGDVLAEYLVPOQGFPCDPAPGAGVNHHRSSSTRSGGDDLTGLPESEE----- 1068
 Db 998 L-----QEALEBEL-----DLDLDLEAEELGATS 1023
 Qy 1069 -----EAPNSPLAPSEB-----AGSDVFDLGMGAAGLQSLPTHD 1105
 Db 1024 LGSALSLPTGTLTRPRGSGSLSPSSGYMPMNSLGEACLDASVILGREGFSPISLH- 1082
 Qy 1106 PPSLQVSEDPYPLPSFTDGYV-----APV-----TC-----SPQPE-----YVNOQPDV 1145
 Db 1083 PLPRGR-----PASESSGHVTSSEAELOEKVSVCRSRSPRPRGDSAYHSQRHS 1135
 Qy 1146 RPQPSPPRSGP-----LPAAAPAGATLERPKTISP-SKNQVY-----KDVFAF 1187
 Db 1136 LPLTPVPLSPGEEBDGNGVWPDTHLRGASSRBCITLSSVGLSSVLTGEEDBD----- 1191
 Qy 1188 GGAIVENPEYLTPOGGAAPORPP 1210
 Db 1192 -----BEVEYNNRRKRGSP-PRPP 1209

RESULT 11

TYFVLV

protein-tyrosine kinase (EC 2.7.1.112) erbB - avian leukosis virus
 N:Contains: amino end of gag protein; env protein fragment; protein-tyrosine kinase
 C:Species: avian leukosis virus, ALV
 C>Date: 31-Dec-1991 #sequence_rev150 31-Dec-1991 #text_change 11-Jun-1999
 C/Accession: B00643; A00643
 R:Nielsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines, M.
 Cell 41, 719-726, 1985
 A>Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pnc
 A/Reference number: A00643; MUID:85228222; PMID:2988784
 A/Accession: B00643
 A/Molecule type: mRNA
 A/Residues: 1-698 <NLT>
 A/Cross-references: GB:M10066; GB:M13881; NID:G211749; PIDN:AAA48763.1; PID:G211750
 A/Note: In Genbank entry CHKBRBF, release 109.0, the source is designated as Gallus gal
 C/Comment: This protein is synthesized as a gag-env-erbB protein.
 C/Genetics:
 A:Gene: gag-env-erbB
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific P
 F.1-6/Product: gag protein (fragment) #status predicted <GAG>
 F.7-59/Product: env protein (fragment) #status predicted <ENV>
 F.60-698/Product: protein-tyrosine kinase erbB #status predicted <ERB>
 F.194-459/Domain: protein kinase homology <KIN>
 F.102-210/Region: protein kinase ATP-binding motif
 F.1229/Active site: Lys #status predicted

Query Match 25.98; Score 1765.5; DB 1; Length 698;
 Best Local Similarity 52.28; Pred. No. 3.5e-65;

Matches 374; Conservative 80; Mismatches 137; Indels 125; Gaps 18;

Qy 578 GPRADOCVACAHYKDPFPCVACPSGVKPDLSYMPKRPDEBACQPCPINCITSVLD 637
 Db 60 GP--DRCMKANRIDSCHCKACRAGVLSGVDTL-VKRYADANAACVCLPNCIRGSKGP 116
 Qy 638 DDKGCPAQRASPLTISAVV-GILLVVLGVVFGILRRQOKIRKYTRRLLOETEL 696

Db 117 GLEGCP---NGSKTSPISAAVGVGGLICLVVVGIGIGLYLRR-HIVKRTLRRLLOEREL 172
 Qy 697 VEPILTPSGAMPQAOQRILKETELRYKVLGSGAFGVYKGIWIPQENYKIPVAIKVL 756
 Db 173 VEPILTPSGAMPQAOQRILKETELRYKVLGSGAFGVYKGIWIPQENYKIPVAIKVL 232
 Qy 757 EHTSPANKELIDEAVYAVAGSPVYSLIGITLSTVOLVLOLMPYGCILDHVRENGR 816
 Db 233 EHTSPANKELIDEAVYAVAGSPVYSLIGITLSTVOLVLOLMPYGCILDHVRENGR 292
 Qy 817 LGSODLLNMOIAKMSVLEDEVRLVHRDLAANVVKSPNHVKITDFGLARLIDIDETE 876
 Db 293 LGSODLLNMOIAKMSVLEDEVRLVHRDLAANVVKSPNHVKITDFGLARLIDIDETE 352
 Qy 877 YEADGKVPIMKMALESILRRRTTHOSDWSYGVTTWELMTFGAKPYDGI PAEIPDLLE 936
 Db 353 YEADGKVPIMKMALESILRRRTTHOSDWSYGVTTWELMTFGAKPYDGI PAEIPDLLE 412
 Qy 937 KEERLPORPITCTIDVYMIWKCMIIDSECPREELVSEFSRMAARDORFVUIONEDLG 995
 Db 413 KEERLPORPITCTIDVYMIWKCMIIDSECPREELVSEFSRMAARDORFVUIONEDLG 472
 Qy 996 PASPLDSTYRSLIEDDDGDLVDAEYLVPOQGFPCDPAPGAGVNHHRSSSTRSG 1055
 Db 473 LPSPTSKPYRTLMESEDMEDIVDAEYLVPOGFP-----NSPT--- 513
 Qy 1056 GEDLTGLPESEEARSPV-----APSEAGSDVFDLGMGAAGLQSLPTHDPSPLQ 1110
 Db 514 -----SRTLSSLSTSNNSANNCID-----RNGQHPVREDSFVQ 550
 Qy 1111 RYSEDPVPLPSET--DGVAAPLTPSPORRYVNOQPDVRRPQPSPPRSGPLPAAPAGATLE 1168
 Db 551 RYSEDPVPLPSET--DGVAAPLTPSPORRYVNOQPDVRRPQPSPPRSGPLPAAPAGATLE 585
 Qy 1169 RPKTISPGKNGVKNQV-----AFGAVENPEYLTPOGGAAPORPPAF 1213
 Db 586 ---TAMVONQIYNNISLTAISGLPMDSKYONSHSTAVDPEYL-----NTNOSPLA 633
 Qy 1214 SPFDNLYYWDQ-----DPEE-----RGAPSTFGKPTAENPEYLGIDVP 1254
 Db 634 KTYFESSPMIGSGNHQINLDNDPYQODPLPNTKXNGLKVAENPEYLRVAAP 689

RESULT 12

TYVUH

protein-tyrosine kinase (EC 2.7.1.112) erbB - avian erythroblastosis virus (strain H)
 C:Species: avian erythroblastosis virus
 C/Date: 18-Apr-1984 #sequence_rev150 18-Apr-1984 #text_change 11-Jun-1999
 C/Accession: A00644; A38022
 R:Yamamoto, T.; Nishida, T.; Miyajima, N.; Kawai, S.; Ooi, T.; Toyoshima, K.
 Cell 35, 71-78, 1983
 A>Title: The erbB gene of avian erythroblastosis virus is a member of the src gene fami
 A/Reference number: A00644; MUID:84026539; PMID:6131329
 A/Molecule type: DNA
 A/Residues: 1-26 'W', 30-119, 'F', 141-145, 'V', 147-152 <DEB>
 A/Cross-references: GB:K02006
 C/Genetics:
 A:Gene: erbB
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C/Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific I
 F.130-395/Domain: protein kinase homology <KIN>
 F.138-146/Region: protein kinase ATP-binding motif
 F.165/Active site: Lys #status predicted

Query March 25, 0%; Score 1702; DB 1; Length 604;
Best Local Similarity 52.2%; Pred. No. 1.2e-62;
Matches 360; Conservative 76; Mismatches 128; Indels 126; Gaps 16;

QY 587 CAHYKDPFCVACRPSGVKPLSYMPIMKPFDEBACCPCEINCHSCVDDDKKCPAEQ 646
DB 3 CAHPIDGHCHCKACAGVLGENDTL-VKRYADANAQCCLPNCRCCKGKGLGECF--- 58
QY 647 RASPLTSTISAVV-GILLVVVLGVVFGILIKRQOKIKRYTMRRLLOETELVEPLTPSGA 705
DB 59 NGSKTPSTIAAGVVGGLLCLVVVGILGILYLR--HIVKRTLRLLQRELRVEPLTPSGE 117
QY 706 MPNDAQWILKETELARKKVLGSGAFGVYGIWTPDEANKIPATVVAEENSPPANK 765
DB 118 APNDAHLKILETEKTEKKVKVLGSGAFGTYGLWTPDEKKIPALKELEAASPPANK 177
QY 766 EILDEAYVMAVGSPYVSRLLGI CLTSTVQVLTQMLPFGCLLDHRENRGRGSGODLLNW 825
DB 178 EILDEAYVMAVDNPHVCRLLGICLTSTVQVLTQMLPFGCLLDYIRKKNIGSQYLLNW 237
QY 826 CMQIAKMSYLEVRLVHRDLAARNVLVKSPNHVKTIDFGARLLDDETEYHADGKVP 885
DB 238 CVQIAKGMNVLSEKRLVHRDLAARNVLVKTPQHYKTIDFGAKLLGDADKEGYHAEKVP 297
QY 886 IKMMALLESILRRRTTHOSDWSYGVTTWELMTFGAKPYDGI-PANEIPDLKEGERLPOPP 945
DB 298 IKMMALLESILHRIYTHOSDWSYGVTTWELMTFGSKPYDGI-PASEISSEVLEKGRLLPOPP 357
QY 946 ICTIDVYVIMVCKMVIDSECRPRPELVSEPSRMAADQRFVVO--NEDLGPASPLOSTP 1004
DB 358 ICTIDVYVIMVCKMVIDSHPKRELVLAESKQARDPRLVIVQGDHMLPSPDTSKF 417
QY 1005 YRSLLDEDMGDLVDAEYELVPOQGFPCDPAPAGAGVHNRHNSSTRSGGDLTLGLE 1064
DB 418 YRTIMEEDMEDVIDADEYELVPHQGF-----NSPST----- 449
QY 1065 PSEBEARPSPL-----APSEAGSDVPFGDGLGMNAKLGSLPTHDSPEPLORYSEDPV 1119
DB 450 -----SRTPLLSLISATSNNSATNCID-----RNGQGHVPRKDSFVQVSGSDPTGN 495
QY 1120 LPSET--DGYAAPLTGSPQPEYVNOQDVPRQPPSPRSEPLPAARPAQTLLRPPTLSRGK 1177
DB 496 PLEESIDDFV-----PAPEYVNO--LMPKKPSTAM----- 524
QY 1178 NGVYKDVFAF-----GGAIVENPEYLPFGGGAAPQHPPPAFSPAFD 1218
DB 525 --VQNGYINPISLTAISKLPMDSRYQNSHSTAVDNPEYL-----NTNQSPLAKTYFE 574
QY 1219 NLTYWDQDPPERGAAPPSTFGKTPAENPEY 1248
DB 575 SSPYMIOSGNHQ-----INLDNDPY 594

RESULT 13
GQPF
epidermal growth factor receptor - fruit fly (Drosophila melanogaster)
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbb
C:Species: Drosophila melanogaster
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 11-Jun-1999
C:Accession: A00640; A38021
R:Lyman, B.; Glaeser, L.; Segal, D.; Schlessinger, J.; Shilo, B.-Z.
A:Title: The Drosophila EGF receptor gene homolog: conservation of both hormone binding
Cell 40, 599-607, 1985
A:Accession: A00640
A:Molecule type: DNA
A:Residues: 1-1330 <LIV>
A:Cross-references: EMBL:K03054
R:Madaworth, S.C.; Vincent III, W.S.; BiloDeau-Wentworth, D.
Nature 314, 178-180, 1985
A:Title: A Drosophila genomic sequence with homology to human epidermal growth factor receptor
A:Reference number: A38021; MUID:85137938; PMID:2983232
A:Accession: A38021

A: Molecule type: DNA
A: Residues: 'A', 832-866, 'V', 866-943, 'QTPSLVK', <NAD>
A: Cross-references: EMBL:X02293; NID:97922; PIDD:CAA26157.1; PID:G929565
C: Comment: This sequence is tentative because the introns have not been identified.
C: Genetics:
A: Gene: FlyBase:Bgr
A: Cross-references: FlyBase:FBgn0003731
A: Map position: 2.57F
C: Superfamily: epidermal growth factor receptor; protein kinase homology
C: Keywords: ATP, autophosphorylation; duplication; glycoprotein; phosphoprotein; phosph
F: 1.732/Domain: extracellular #status predicted <EXT>
F: 723-764/Domain: transmembrane #status predicted <TM>
F: 765-1330/Domain: intracellular #status predicted <INT>
F: 808-1072/Domain: protein kinase homology <KIN>
F: 816-824/Region: protein kinase ATP-binding motif
F: 122,300,324,363,518,668,695,700/Binding site: carbohydrate (Asn) (covalent) #status p
F: 172/Binding site: phosphatase (Thr) (covalent) (by protein kinase C) #status predicted
F: 843/Active site: Lys #status predicted
F: 1181/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predic

Query Match 24.3%; Score 1653.5; DB 1; Length 1330;
Best Local Similarity 30.0%; Pred. No. 2.4e-60;
Matches 414; Conservative 179; Mismatches 415; Indels 371; Gaps 39;

Qy	80	VQGVTLIAHQVROVLPQRLRIYRGTOLF----	EDNYALAVLDNGDLNNTPTVGTASP	134
Db	38	ITNYIVIGLIDLPCTLSYRLQIRGRLFSIYSEKXALFV-----	TY	81
Qy	135	GGRLRLRLRLTEILKSGVLIQRNPOLCYODTIMKDI	PHNQNGLALTLIDTNSRACHP	194
Db	82	SKATLTLEIPDLRDVLNQGVGFHNHNNYLCHMKTIQWSEIVSGTAYNYDFTAPERECPK	141	
Qy	195	CSPMCKSRRCMGSSSDCOSLTRYVCAGGA--RCKGPLPDCCHBOCAACTGPKHSDC	252	
Db	142	CHSECTHG--CMGSEPKKCKQKFSKLTCSPOCAGKCYCPKRECHLFCAGCTGTGTQDC	200	
Qy	253	LACLHPHSGICELHCPALVYNTDTPESMPNPEGRYTFGASCYTACPYNLTSDVSGCT	312	
Db	201	IACKNPFDEAVSKSECPKPKRYNPTTYLETNPEGKAYAGATCYVECP--GHILRDNGACV	259	
Qy	313	LVNCLAHNGEVLAEIDGTORCKCKSKPCARVCYGLGMEHLREVRATVSANIQPACCKIFG	372	
Db	260	RSCQDQKMDGGE-----CVPCKGPKCTCPGVVLH-----	AGIDSFNCTVIDG	306
Qy	373	SLAFLPESFDG--DPASNTA-----	PLQEQLOVFETLEETGYLYTSANPDSLPLDSLV	424
Db	307	NIRLDOTFSGPDVANYMTGMRYIPLDEBRREVPSTVAKEITGYLNEIGHHPQRLNLSY	366	
Qy	425	FQNLQVIRGRILHNGAY--SLTLOGLISWLGSLRLSELGSLALIHNTHLCPVHTVWD	483	
Db	367	FRNLLETIHGRQLMBSMPALALIVKSLYSLEMRYLKQISSGVVIOHNRDLCYVSNIRMP	426	
Qy	484	QLPKNPQALLHTRANRDEDC-----		504
Db	427	ALQEPQKQWVNNENLRADLGKFLTLLISVQHNIIHIFAI	CREKNHLLGSVQGRLL	486
Qy	505	-----		504
Db	487	GSNHSQVPLYQELDFQMHRLRLMLYIQVGLINSIQDSNKHQLTDACYSVPVPSLTIER	546	
Qy	505	-----		510
Db	547	AFVAYIAGLAMELQITARSASWRHSTKLPAEGQVPRWVFLGVCSARAGIAEPLAGR	606	
Qy	511	-----		562
Db	607	AVCRCKHPRLCELCTNNGYHNOVCSKTHYRRRBOCTREC-----	PADHYTDEBORBCFQ	660
Qy	563	CHPECKQNGSVTCFGEADOCVAHYK-----	DPF-----	608
Db	661	RHPEC--NG-----	CTGEGADCKSCRNFKLPDANETQPYNSTMFNCTSKPPLMRHNY	714
Qy	609	SYNPWKPDPDEGACQCPINCTHS	CVDDLDDKGPABQARSPLTISI	668


```

Db      715 QYALGPY-----CAASPRSSKLTAND-----VNMIFITGAVLFTIC 755
Qy      669 VVFGI-LIKRROOKIRKRYT--MRLLLOETELVEPLTPSGAMPQAOIRILKETELRKYV 725
Db      756 ILCVYIYICRQKAKKERTYKMTMALSGRDSERLPRNSIGALCKLRIYKDAELKRGV 815
Qy      726 LSGAGFTYVYKGIWIDGENVKI PVAIKVLRNTSPANKREILDEAYVMAGVSPYVSR 785
Db      816 LGMGAGRYKGVWVEGENVKI PVAIKELKSTGASSSEFLREAVIMASBEHVNLKL 875
Qy      786 LGLCLSTVQVQLMPYGLLDHYRNRGRSLGSDLLANCMQIAGMSTLEDVRLYHRD 845
Db      876 LVMGSSQWMLITQLMPLGCLLDYVNRNRDKISKLLNMSTQIAKMSYLEERLVRHD 935
Qy      846 LAARNVLYK---SPNVKITDPIGLARLLIDETEYHADGKVPVKMMALSLIRRPETHQ 902
Db      936 LAARNVLYKLAGEDH---DFGLAKLSSDSNRYKAGAKMFKMLALECIRNRVPTSK 991
Qy      903 SDVMSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGERLPQPICTIDVYIMVKCMMD 962
Db      992 SDVMAGVTIMELTGTGGRHENIPADIPDLIEVGLKEQPEICSDIYCTILSCWHD 1051
Qy      963 SECRPFRELVSFPMARDPQRFVYIQNEDLG--PASPLDSTPYSLLED--DMGL 1017
Db      1052 AAMRPFPKQITVFAFARPRGYLAIGDKFTRLPA-----YTSQDEKDIRKLAPT 1104
Qy      1018 VAABEVLVPOQGFPCDPAPGAGMYHRRSSSTRSGGDLTLGLEPSEEARP----- 1071
Db      1105 TGQSEAIAPBDYLOKRALGPS-----HRTDCT-----DENPKLNRVC 1143
Qy      1072 RSPPLASSEGASDVDFG--DLGMAKAGLSLPTHDPBLORYSEDPVPLPSETDGYV 1128
Db      1144 KQPSNKGSTGDERDSAREVGVNLR-----LDLPVEDDDYL 1182
Qy      1129 APLTGPQPEYVQDVRPQPPSPREGPLPAAPPAATLERPRLTSGKGVKQDVAFG 1188
Db      1183 MP-TCPGPPNNNNNM-----NPNQNMMAVGAAGYM-----DLIGVP 1220
Qy      1189 GAVENDEYL-----TPGGAAPOPH-----PPAFSP-APDNLYWMD 1224
Db      1221 VSDNREYLNAQTLGVGESPIPTQITIGVWGPGPMGVKVPMPGSEPTSDHEIYND 1279

RESULT 14
S35745
protein-tyrosine kinase (EC 2.7.1.112) erbB - avian erythroblastosis virus
C/Species: avian erythroblastosis virus
C/Date: 03-Mar-1994 #sequence_revision 26-May-1995 #text_change 28-Feb-1997
C/Accession: S35745
R/Vennstrom, B.
submitted to the EMBL Data Library, March 1993
A/Reference number: S35743
A/Molecule type: DNA
A/Residues: 1-544 <VEN>
A/Cross-references: EMBL:X12707
C/Genetics:
A/Gene: erbB
C/Superfamily: epidermal growth factor receptor; protein kinase homology
C/Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific
P/135-400/Domain: protein kinase homology <KIN>
P/143-151/Region: protein kinase ATP-binding motif
P/170/Active site: Lys #status predicted

```

```

Query Match      24.1%; Score 1645; DB 2; Length 544;
Best Local Similarity 57.7%; Pred. No. 2.3e-60;
Matches 335; Conservative 65; Mismatches 97; Indels 84; Gaps 13;

Qy      578 GPADQVCAAHKDPFCVACRPSGVKPLSTYMPIRKPPDEBACQPCPINTCHSCVDL 637
Db      1 GP--DHCMKCAHPIDEPHCVKACPAVGLGENDTL-VKRYADANAVCQLCHPNCIRGCKGP 57

```

```

Qy      638 DPKCPAEQASPLTIIISAV-GILLVVVLGVVFGILIRROOKIRKTYMRLLQETEL 696
Db      58 GLEGCP-----NGSKTPIAAGVVGGLCLVVGIGIGLYLRR-HIVKRTLRLLOEREL 113
Qy      697 VEPLTPSGAMPQAOIRILKETELRKYVTSGAFGVYVYIMPDENVKIPVAIKYLR 756
Db      114 VEPLTPSGAMPQAOIRILKETELRKYVTSGAFGVYVYIMPDENVKIPVAIKYLR 173
Qy      757 ENTPSPANKREILDEAYVMAGVSPYVSRLLGICLTSTVQLVTOQLMPYGLLDHYRNRGR 816
Db      174 EATSPANKREILDEAYVMSVDNPHVCRLLGICLTSTVQLVTOQLMPYGLLDHYRNRGR 233
Qy      817 LGSODLLANCMQIAKMSYLEDVRLYHRDIAANVLYKSPNHYKITDPGLARLLIDETE 876
Db      234 LGSQVLLNMCVQIAKMSYLEDVRLYHRDIAANVLYKSPNHYKITDPGLARLLIDETE 293
Qy      877 YEADGKVPVKMMALSLIRRPETHQSDVMSYGVTVWELMTFGAKPYDGI PAEIPDLLE 936
Db      294 YEADGKVPVKMMALSLIRRPETHQSDVMSYGVTVWELMTFGAKPYDGI PAEIPDLLE 353
Qy      937 KGERLPQPICTIDVYIMVKCMMDSECRPFRELVSFPMARDPQRFVYIQNEDLG 995
Db      354 KGERLPQPICTIDVYIMVKCMMDSDASRPFRELVAESKQARDPRLVIGDERMH 413
Qy      996 PASPLDSTPYSLLEDMDGLVDABEYLVPOQGFPCDPAPGAGMYHRRSSSTRSG 1055
Db      414 LPSPTSKFYRLTMEEDMEDIVDADEYLVPHQGF-----NSPST--- 454
Qy      1056 GBDLTLGLEPSEEARSP-----APSEGASDVDFGDLGMAKAGLSLPTHDPSPLO 1110
Db      455 -----SRPLSLSLATSNNSATNCDIRNG-----H----- 481
Qy      1111 RYSEDPVPLPSETDGYVAPLTCSPOPEYVQDVRPQPS 1151
Db      482 -----PVAREDFL-----PAPEYVQ--LMPKKS 504

```

```

RESULT 15
S00727
kinase-related transforming protein (erbB) (EC 2.7.1.1) - avian erythroblastosis virus
C/Species: avian erythroblastosis virus
C/Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 28-Feb-1997
C/Accession: S00727
R/Scotting, P.; Vennstrom, B.; Jansen, M.; Graf, T.; Beug, H.; Hayman, M.J.
Oncogene Res. 1, 265-278, 1987
A/Title: Common site of mutation in the erbB gene of avian erythroblastosis virus mutan
A/Reference number: S00727; MID:88217326; PMID:2897102
A/Accession: S00727
A/Molecule type: DNA
A/Residues: 1-545 <SCO>
A/Cross-references: EMBL:X06943
C/Genetics:
A/Gene: erbB
C/Superfamily: epidermal growth factor receptor; protein kinase homology
C/Keywords: ATP; phosphotransferase
P/135-400/Domain: protein kinase homology <KIN>
P/143-151/Region: protein kinase ATP-binding motif

```

```

Query Match      24.0%; Score 1638; DB 2; Length 545;
Best Local Similarity 57.7%; Pred. No. 4.4e-60;
Matches 335; Conservative 64; Mismatches 98; Indels 84; Gaps 13;

Qy      578 GPADQVCAAHKDPFCVACRPSGVKPLSTYMPIRKPPDEBACQPCPINTCHSCVDL 637
Db      1 GP--DHCMKCAHPIDEPHCVKACPAVGLGENDTL-VKRYADANAVCQLCHPNCIRGCKGP 57
Qy      638 DPKCPAEQASPLTIIISAV-GILLVVVLGVVFGILIRROOKIRKTYMRLLQETEL 696
Db      58 GLEGCP-----NGSKTPIAAGVVGGLCLVVGIGIGLYLRR-HIVKRTLRLLOEREL 113
Qy      697 VEPLTPSGAMPQAOIRILKETELRKYVTSGAFGVYVYIMPDENVKIPVAIKYLR 756
Db      114 VEPLTPSGAMPQAOIRILKETELRKYVTSGAFGVYVYIMPDENVKIPVAIKYLR 173

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2004, 16:06:24 ; Search time 18 Seconds
(without alignments)
3278.806 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815
Sequence: 1 MELALCRWGLLALLPPGA.....TFKGTPTANPEYLGLDVPV 1255

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	1 ERB2_HUMAN	P04626 homo sapien
2	6003	88.1	1257	1 ERB2_RAT	P06494 rattus norv
3	5993.5	87.9	1254	1 ERB2_MESAU	Q60833 mesocricetu
4	3167	46.5	1210	1 EGFR_HUMAN	P00533 homo sapien
5	3146	46.2	1210	1 EGFR_MOUSE	Q01279 mus musculu
6	3004.5	44.1	1308	1 ERB4_HUMAN	Q15303 homo sapien
7	2989	43.9	1308	1 ERB4_RAT	Q62856 rattus norv
8	2724.5	40.0	1167	1 XMRK_XIPMA	P13388 xiphophorus
9	2441.5	35.8	1339	1 ERB3_HUMAN	P13388 homo sapien
10	2370.5	34.8	1339	1 ERB3_RAT	Q62799 rattus norv
11	1975	29.0	1426	1 EGFR_DROME	P04412 drosophila
12	1748.5	25.7	634	1 ERBB_ALV	P00534 avian leuko
13	1702	25.0	604	1 ERBB_AYIER	P00535 avian eryth
14	1628	23.9	540	1 ERBB_AITIU	P11373 avian eryth
15	1623	23.8	703	1 EGFR_CHICK	P13387 gallus gall
16	1301	19.1	1367	1 LTR3_CABRL	P24348 caenorhabdi
17	1142.5	16.8	245	1 ERB2_MOUSE	P70424 mus musculu
18	736	10.8	1363	1 ILPR_BRALA	O02466 brachiocto
19	715	10.5	1382	1 INSR_HUMAN	P06213 homo sapien
20	710	10.4	1383	1 INSR_RAT	P15127 rattus norv
21	709.5	10.4	1372	1 INSR_MOUSE	P15208 mus musculu
22	706	10.4	1300	1 IRR_MOUSE	Q9W140 mus musculu
23	703.5	10.3	1607	1 MIRP_LYMER	Q25410 lymphoma ara
24	699	10.3	1297	1 IRR_HUMAN	P14616 homo sapien
25	694.5	10.2	1300	1 IRR_CAVPO	P14617 cavia porce
26	682.5	10.0	1477	1 HTR7_HYDAT	Q25197 hydra atten
27	651	9.6	1367	1 EGFR_HUMAN	P00869 homo sapien
28	643	9.4	1373	1 IGR_MOUSE	Q60751 mus musculu
29	639.5	9.4	1370	1 IGR_RAT	P24062 rattus norv
30	627	9.2	1390	1 INSR_AEDAE	Q93105 aedes aegypt
31	622	9.1	2146	1 INSR_DROME	P09208 drosophila
32	607	8.9	987	1 EPB4_HUMAN	P54760 homo sapien
33	590.5	8.7	984	1 EPB1_CHICK	O07944 gallus gall

34	589.5	8.7	977	1 EPB2_MOUSE	Q03145 mus musculu
35	587	8.6	1114	1 RER_HUMAN	P07949 homo sapien
36	585	8.6	987	1 EPB4_MOUSE	P54761 mus musculu
37	584.5	8.6	976	1 EPB2_HUMAN	P29317 homo sapien
38	582.5	8.5	984	1 EPB1_RAT	P09759 rattus norv
39	578	8.5	902	1 EPBB_XENLA	Q91736 xenopus lae
40	576.5	8.5	984	1 EPB1_HUMAN	P54762 homo sapien
41	576.5	8.5	985	1 EPBA_XENLA	Q91571 xenopus lae
42	575	8.4	1053	1 FAK1_CHICK	Q00944 gallus gall
43	569	8.3	1068	1 FAK1_XENLA	Q91738 xenopus lae
44	567	8.3	757	1 HTR6_HYDAT	P53356 hydra atten
45	566	8.3	1052	1 FAK1_HUMAN	Q05397 homo sapien

ALIGNMENTS

RESULT 1	ERB2_HUMAN	STANDARD;	PRT; 1255 AA.
AC	P04626;		
AD	13-AUG-1987 (Rel. 05, Created)		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	13-SEP-2003 (Rel. 42, Last annotation update)		
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)		
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell		
DE	surface receptor HER2) (MLN 19).		
GN	ERB2 OR HER2 OR NGL OR NEU.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_Taxid9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86118663; PubMed=3003577;		
RA	Yamamoto T., Ikawa S., Akiyama T., Sema K., Nomura N., Miyajima N.,		
RA	Saito T., Toyoshima K.;		
RT	"Similarity of protein encoded by the human c-erb-B-2 gene to		
RT	epidermal growth factor receptor.";		
RL	Nature 319:230-234 (1986).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86070181; PubMed=2999974;		
RA	Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,		
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,		
RA	Frankel U., Levinson A., Ulrich A., Schlesinger J.,		
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor		
RT	shares chromosomal location with neu oncogene.";		
RL	Science 230:1132-1139 (1985).		
RN	[3]		
RP	SEQUENCE OF 737-1031 FROM N.A.		
RX	MEDLINE=86016729; PubMed=2995967;		
RA	Sema K., Kamata N., Toyoshima K., Yamamoto T.,		
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the		
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a		
RT	human salivary gland adenocarcinoma.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501 (1985).		
RN	[4]		
RP	VARIANTS VAL-654 AND VAL-655.		
RX	MEDLINE=93194196; PubMed=8095488;		
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.,		
RT	"Characterization of a new allele of the human ERB2 gene by allele-		
RT	specific competition hybridization.";		
RL	Genomics 15:426-429 (1993).		
CC	- FUNCTION: Essential component of a neurotrophin-receptor complex,		
CC	although neurotrophins do not interact with it alone. GP30 is a		
CC	potential ligand for this receptor. Not activated by EGF, TGF-		
CC	alpha and amphiregulin.		
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein		
CC	tyrosine phosphate.		
CC	- SUBUNIT: Heterodimer with each of the other ERBB receptors		
CC	(Potential). Interacts with PRKCAP (By similarity).		
CC	- SUBCELLULAR LOCATION: Type I membrane protein.		

CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine
 CC residues (By similarity).
 CC -1- POLYMORPHISM: There are four alleles due to the variations in
 CC positions 654 and 655. Allele B1 (654-ile-ile-655) has a frequency
 CC of 0.782; allele B2 (654-ile-val-655) has a frequency of 0.206;
 CC allele B3 (654-val-val-655) has a frequency of 0.012.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M11767, AAA35808.1, -
 CC EMBL, M11761, AAA35808.1, JOINED.
 CC EMBL, M11762, AAA35808.1, JOINED.
 CC EMBL, M11763, AAA35808.1, JOINED.
 CC EMBL, M11764, AAA35808.1, JOINED.
 CC EMBL, M11765, AAA35808.1, JOINED.
 CC EMBL, M11766, AAA35808.1, JOINED.
 CC EMBL, M11730, AAA35808.1, JOINED.
 CC EMBL, M12036, AAA35978.1, -
 CC PIR, A24571, A24571.
 CC PIR, A24571, A24571.
 CC PDB, 1N8Z, 16-FEB-03.
 CC PDB, 1OR1, 01-JAN-00.
 CC GeneW, HGNC:3430, ERBB2.
 CC MIM, 164870, -
 CC DR GO:0005012, F:New/Erbb-2 receptor activity; TMS.
 CC DR GO:0004716, F:Receptor signaling protein tyrosine kinase . . . ; TMS.
 CC DR GO:0008283, P:cell proliferation; TMS.
 CC DR GO:0007048, P:oncogenesis; TMS.
 CC DR GO:0006470, P:protein amino acid dephosphorylation; TMS.
 CC DR GO:0006466, P:protein amino acid phosphorylation; TMS.
 CC DR InterPro: IPR000494, EGFR L domain.
 CC DR InterPro: IPR006212, Furin-like.
 CC DR InterPro: IPR00719, Prot. repeat.
 CC DR InterPro: IPR001245, Tyr. kinase.
 CC DR InterPro: IPR004019, YLP motif.
 CC DR Pfam: PF00757, Furin-like; 1.
 CC DR Pfam: PF00069, Pkinase; 1.
 CC DR Pfam: PF01030, Recep_L_domain; 2.
 CC DR Pfam: PF02757, YLP; 2.
 CC DR PRINTS: PR00109, TYRKINASE.
 CC DR PRODOM: PD000001, Prot_kinase; 1.
 CC DR SMART: SM00261, FU; 4.
 CC DR SMART: SM00219, TyrcK; 1.
 CC DR PROSITE: PS00107, PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE: PS00109, PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE: PS50011, PROTEIN_KINASE_DOM; 1.
 CC DR Transmembrane, Glycoprotein, Multigene family; Receptor; Signal;
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 CC Polymorphism; 3D-structure.
 CC FT SIGNAL 1 21 POTENTIAL.
 CC FT CHAIN 22 125 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 CC FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 653 675 POTENTIAL.
 CC FT DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).
 CC FT NP BIND 720 987 PROTEIN KINASE.
 CC FT BINDING 726 734 ATP (BY SIMILARITY).
 CC FT BINDING 753 753 ATP (BY SIMILARITY).
 CC FT ACT_SITE 845 845 BY SIMILARITY.
 CC FT DISULFID 195 204 BY SIMILARITY.
 CC FT DISULFID 199 212 BY SIMILARITY.
 CC FT DISULFID 220 237 BY SIMILARITY.
 CC FT DISULFID 234 235 BY SIMILARITY.
 CC FT DISULFID 236 244 BY SIMILARITY.
 CC FT DISULFID 240 252 BY SIMILARITY.
 CC FT DISULFID 255 264 BY SIMILARITY.

FT DISULFID 268 295 BY SIMILARITY.
 FT DISULFID 299 311 BY SIMILARITY.
 FT DISULFID 315 331 BY SIMILARITY.
 FT DISULFID 334 338 BY SIMILARITY.
 FT DISULFID 511 520 BY SIMILARITY.
 FT DISULFID 515 528 BY SIMILARITY.
 FT DISULFID 531 540 BY SIMILARITY.
 FT DISULFID 544 560 BY SIMILARITY.
 FT DISULFID 563 576 BY SIMILARITY.
 FT DISULFID 567 584 BY SIMILARITY.
 FT DISULFID 587 596 BY SIMILARITY.
 FT DISULFID 600 623 BY SIMILARITY.
 FT DISULFID 626 634 BY SIMILARITY.
 FT DISULFID 630 642 BY SIMILARITY.
 FT MOD_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 654 654 I -> V (IN DBSNP:1801201).
 FT VARIANT 655 655 I -> V (IN DBSNP:1801200).
 FT VARIANT 655 655 /FTID=VAR_004077.
 FT CONFLICT 1170 1170 P -> A (IN REF. 2).
 FT SEQUENCE 1255 AA; 137909 MW; 3959DFDA04DCFP62 CRC64;
 SQ
 Query Match 100.0%; Score 6815; DB 1; Length 1255;
 . Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MELALCRWGLLLALPPGASTVCTGTDKRLRPAPEHLDMLRHYOCQVQVQGL 60
 1 MELALCRWGLLLALPPGASTVCTGTDKRLRPAPEHLDMLRHYOCQVQVQGL 60
 61 ELTYLPTNASLFLDIOEVOGVYLIANOVQVPLQRLRIYRGTLFEDNYALVLDNG 120
 61 ELTYLPTNASLFLDIOEVOGVYLIANOVQVPLQRLRIYRGTLFEDNYALVLDNG 120
 61 ELTYLPTNASLFLDIOEVOGVYLIANOVQVPLQRLRIYRGTLFEDNYALVLDNG 120
 121 DLINNTTPTVGTASPGARELQRLSTLTKGGLVLIQNNPQLCYDTIIMKDIFFHKNOLA 180
 121 DLINNTTPTVGTASPGARELQRLSTLTKGGLVLIQNNPQLCYDTIIMKDIFFHKNOLA 180
 121 DLINNTTPTVGTASPGARELQRLSTLTKGGLVLIQNNPQLCYDTIIMKDIFFHKNOLA 180
 121 DLINNTTPTVGTASPGARELQRLSTLTKGGLVLIQNNPQLCYDTIIMKDIFFHKNOLA 180
 121 DLINNTTPTVGTASPGARELQRLSTLTKGGLVLIQNNPQLCYDTIIMKDIFFHKNOLA 180
 181 LTLIDTNRSRACHPCSPCKSRGSESSDQSLTRTVCAAGCARCKGPLPTDCHEQC 240
 181 LTLIDTNRSRACHPCSPCKSRGSESSDQSLTRTVCAAGCARCKGPLPTDCHEQC 240
 181 LTLIDTNRSRACHPCSPCKSRGSESSDQSLTRTVCAAGCARCKGPLPTDCHEQC 240
 181 LTLIDTNRSRACHPCSPCKSRGSESSDQSLTRTVCAAGCARCKGPLPTDCHEQC 240
 181 LTLIDTNRSRACHPCSPCKSRGSESSDQSLTRTVCAAGCARCKGPLPTDCHEQC 240
 241 AAGCTGPRHSCLACIHFNHSIGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACP 300
 241 AAGCTGPRHSCLACIHFNHSIGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACP 300
 241 AAGCTGPRHSCLACIHFNHSIGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACP 300
 241 AAGCTGPRHSCLACIHFNHSIGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACP 300
 241 AAGCTGPRHSCLACIHFNHSIGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACP 300
 301 YNYLSTDVGSCTLVCPILNNOEVTADGTORCEKSKPCARVCYGLGMEHLREAVTASN 360
 301 YNYLSTDVGSCTLVCPILNNOEVTADGTORCEKSKPCARVCYGLGMEHLREAVTASN 360
 301 YNYLSTDVGSCTLVCPILNNOEVTADGTORCEKSKPCARVCYGLGMEHLREAVTASN 360
 301 YNYLSTDVGSCTLVCPILNNOEVTADGTORCEKSKPCARVCYGLGMEHLREAVTASN 360
 301 YNYLSTDVGSCTLVCPILNNOEVTADGTORCEKSKPCARVCYGLGMEHLREAVTASN 360
 361 IOEPAGCKKIGSLAFLESFDDGDPASNTAPLOPOLOVFTLEITGLYISAMPDSL 420
 361 IOEPAGCKKIGSLAFLESFDDGDPASNTAPLOPOLOVFTLEITGLYISAMPDSL 420
 361 IOEPAGCKKIGSLAFLESFDDGDPASNTAPLOPOLOVFTLEITGLYISAMPDSL 420
 361 IOEPAGCKKIGSLAFLESFDDGDPASNTAPLOPOLOVFTLEITGLYISAMPDSL 420
 361 IOEPAGCKKIGSLAFLESFDDGDPASNTAPLOPOLOVFTLEITGLYISAMPDSL 420
 421 DIISVQNTQVIRGLIHNGAYSLTQGLIGISWLGRLSRELSGGLALIHNNHLCFVHTV 480
 421 DIISVQNTQVIRGLIHNGAYSLTQGLIGISWLGRLSRELSGGLALIHNNHLCFVHTV 480
 421 DIISVQNTQVIRGLIHNGAYSLTQGLIGISWLGRLSRELSGGLALIHNNHLCFVHTV 480
 421 DIISVQNTQVIRGLIHNGAYSLTQGLIGISWLGRLSRELSGGLALIHNNHLCFVHTV 480
 421 DIISVQNTQVIRGLIHNGAYSLTQGLIGISWLGRLSRELSGGLALIHNNHLCFVHTV 480
 481 PWDOLFERNPHALLHTANRPEDECVGEGGLACHQLCARGHGWPPTQVNCQFRLGQRC 540
 481 PWDOLFERNPHALLHTANRPEDECVGEGGLACHQLCARGHGWPPTQVNCQFRLGQRC 540
 481 PWDOLFERNPHALLHTANRPEDECVGEGGLACHQLCARGHGWPPTQVNCQFRLGQRC 540
 481 PWDOLFERNPHALLHTANRPEDECVGEGGLACHQLCARGHGWPPTQVNCQFRLGQRC 540
 481 PWDOLFERNPHALLHTANRPEDECVGEGGLACHQLCARGHGWPPTQVNCQFRLGQRC 540
 541 VEECVLQGLPREYVNAHCHLPCHCEQOPONSVTCPFEADQCVACAHYKDPFCVAC 600
 541 VEECVLQGLPREYVNAHCHLPCHCEQOPONSVTCPFEADQCVACAHYKDPFCVAC 600
 541 VEECVLQGLPREYVNAHCHLPCHCEQOPONSVTCPFEADQCVACAHYKDPFCVAC 600
 541 VEECVLQGLPREYVNAHCHLPCHCEQOPONSVTCPFEADQCVACAHYKDPFCVAC 600
 541 VEECVLQGLPREYVNAHCHLPCHCEQOPONSVTCPFEADQCVACAHYKDPFCVAC 600

QY 601 PSQVPSDLSYPIWKPDEBEGACQPCPINCTHSCVDLDDKCGPAQORASPLTISIAVVG 660
 DB 601 PSQVPSDLSYPIWKPDEBEGACQPCPINCTHSCVDLDDKCGPAQORASPLTISIAVVG 660
 QY 661 ILLVVVGLVVFGLILKRRQOKIRKTYMRLLQETELVPLTSSGAMPQOAKRILKEVEL 720
 DB 661 ILLVVVGLVVFGLILKRRQOKIRKTYMRLLQETELVPLTSSGAMPQOAKRILKEVEL 720
 QY 721 RKVKVLSGAGFVYKGIWIPDGENVKIPIAKVLRENTSPIANKBIIDEAYVMAGVSP 780
 DB 721 RKVKVLSGAGFVYKGIWIPDGENVKIPIAKVLRENTSPIANKBIIDEAYVMAGVSP 780
 QY 781 YVSRLLGICLTSTVOLVTLMPYGGCLDHRVRENRGLSGQDLNMCQIAKMSYLEIDVR 840
 DB 781 YVSRLLGICLTSTVOLVTLMPYGGCLDHRVRENRGLSGQDLNMCQIAKMSYLEIDVR 840
 QY 841 LVHDDLAAARNLVKSPNNVKTIDBGLARLLDIDETEHYADGKVPITKMMALSLRRRT 900
 DB 841 LVHDDLAAARNLVKSPNNVKTIDBGLARLLDIDETEHYADGKVPITKMMALSLRRRT 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYMIWKKM 960
 DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYMIWKKM 960
 QY 961 IDSECRPRFRELVSFSSWARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
 DB 961 IDSECRPRFRELVSFSSWARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
 QY 1021 BEYLVPOGFCFCDPAFAGAGMWHHRHSSSTRSGGDLTLGLBSEBAPSPPLAPSEB 1080
 DB 1021 BEYLVPOGFCFCDPAFAGAGMWHHRHSSSTRSGGDLTLGLBSEBAPSPPLAPSEB 1080
 QY 1081 AGSDVFDGDLGMAKGLQSLPTHDSPLOKYSBPTVPLPSETGYVAPLCPQPEYV 1140
 DB 1081 AGSDVFDGDLGMAKGLQSLPTHDSPLOKYSBPTVPLPSETGYVAPLCPQPEYV 1140
 QY 1141 NQPDVRFQPSPRBEGPLPAARPAATLERPKTLSPKNGVVDVAFGAVENPEYLPQ 1200
 DB 1141 NQPDVRFQPSPRBEGPLPAARPAATLERPKTLSPKNGVVDVAFGAVENPEYLPQ 1200
 QY 1201 CGAAPQHPHPPAFSPFNLVWDDPPBRKAPSTFKTPTAENPEYLGDLVPV 1255
 DB 1201 CGAAPQHPHPPAFSPFNLVWDDPPBRKAPSTFKTPTAENPEYLGDLVPV 1255

RA Lat C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially
 RL expressed in the vertebrate nervous system";
 RN Neuron 6:691-704(1991).
 RP STRUCTURE BY NMR OF 650-668.
 RX MEDLINE=92155181; PubMed=1346763;
 RA Gullick W.J., Bottomley A.C., Lofes F.J., Doak D.G., Mulvey D.,
 RT Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 "Three dimensional structure of the transmembrane region of the proto-
 oncogenic and oncogenic forms of the new protein";
 RL EMBL J.11:43-48(1992).
 CC -1- FUNCTION: Essential component of a neurogulin-receptor complex,
 although neurogulins do not interact with it alone. GP30 is a
 potential ligand for this receptor. Not activated by EGF, TGF-
 alpha and amphiregulin.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 tyrosine phosphate.
 CC -1- SUBUNIT: Heterodimer with each of the other ERBB receptors. The
 constitutively activated oncogenic variant forms a homodimer.
 CC Interacts with PRKCAP (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine
 residues (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL, X03362; CAA27059.1; ALF_INIT.
 DR PDB: 1ITJ; 27-JUN-01.
 DR PDB: 1NB7; 18-FEB-03.
 DR InterPro: IPR000494; EGFR_L domain.
 DR InterPro: IPR006211; Furin-like.
 DR InterPro: IPR006212; Furin repeat.
 DR InterPro: IPR000719; Prot_Kinase.
 DR InterPro: IPR001245; Tyr_Kinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; Kinase; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR Pfam: PF02757; YLP_2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_Kinase; 1.
 DR SMART; SMO0261; FU; 4.
 DR SMART; SMO0219; TyKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Proto-oncogene; Disease mutation; 3D-structure.
 FT CHAIN 1..21
 FT DOMAIN 22..1257 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT TRANSMEM 655..677 POTENTIAL.
 FT DOMAIN 678..1257 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 159..369 CYS-RICH.
 FT DOMAIN 473..646 CYS-RICH.
 FT DOMAIN 722..989 PROTEIN KINASE.
 FT NP_BIND 728..736 ATP (BY SIMILARITY).
 FT BINDING 755..755 ATP (BY SIMILARITY).
 FT ACT_SITE 847..847 BY SIMILARITY.
 FT DISULFID 196..205 BY SIMILARITY.
 FT DISULFID 200..213 BY SIMILARITY.
 FT DISULFID 221..228 BY SIMILARITY.
 FT DISULFID 225..236 BY SIMILARITY.
 FT DISULFID 237..245 BY SIMILARITY.

FT DISULFID 241 253 BY SIMILARITY.
 FT DISULFID 256 265 BY SIMILARITY.
 FT DISULFID 269 296 BY SIMILARITY.
 FT DISULFID 300 312 BY SIMILARITY.
 FT DISULFID 316 332 BY SIMILARITY.
 FT DISULFID 335 339 BY SIMILARITY.
 FT DISULFID 513 522 BY SIMILARITY.
 FT DISULFID 517 530 BY SIMILARITY.
 FT DISULFID 533 542 BY SIMILARITY.
 FT DISULFID 546 562 BY SIMILARITY.
 FT DISULFID 565 578 BY SIMILARITY.
 FT DISULFID 569 586 BY SIMILARITY.
 FT DISULFID 589 598 BY SIMILARITY.
 FT DISULFID 602 625 BY SIMILARITY.
 FT DISULFID 628 636 BY SIMILARITY.
 FT DISULFID 632 644 BY SIMILARITY.
 FT MOD_RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1250 1250 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).
 SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;

Query Match 88.1%; Score 6003; DB 1; Length 1257;
 Best Local Similarity 88.0%; Pred. No. 7, 6e-310;

Matches 1106; Conservative 49; Mismatches 100; Indels 2; Gaps 2;

QY 1 MELALCRMGLTALLPFGASTOVCTGTDMRLPASPETHLDMRLHLYGCCQVVGNTL 60
 DB 1 MELAMCMGFLTALLPFGIAGTQCTGTDMRLPASPETHLDMRLHLYGCCQVVGNTL 60
 QY 61 ELTTLPTNASSFLDDIOEVQSVTIANQVQVPLQRLRYRGTQLPFDNTALAVLNG 120
 DB 61 ELTYVPAVASLFLDDIOEVQSVTIANQVQVPLQRLRYRGTQLPFDNTALAVLNR 120
 QY 121 DPLNNTTFTV- GASPGSLREIOLRSITELIKGGLVLIORPOLCYDDTIIMKDI FHKNNOL 179
 DB 121 DPODVAASTPRTYRGLRELOLRSLTELKGGVLRGNPOLCYDDTIIMKDI FHKNNOL 180
 QY 180 ALLTIDNRSRACHP CSPMKCSRCMGSSESDCSLTITVCAAGGACRCKPLPTDCHEQ 239
 DB 181 APVDIDNRSRACHP CSPMKCSRCMGSSESDCSLTITVCAAGGACRCKPLPTDCHEQ 240
 QY 240 CAAGCTGPRGSDCLALCPHNSGICELHCPALVTNTDTFESMPNREGRTTFGASCVTAC 299
 DB 241 CAAGCTGPRGSDCLALCPHNSGICELHCPALVTNTDTFESMPNREGRTTFGASCVTTC 300
 QY 300 PNYLSTVGSGCTIACPLNDEVTABDGTORCEKSKPCARCYGIGMHLREVAVISA 359
 DB 301 PNYLSTVGSGCTIACPLNDEVTABDGTORCEKSKPCARCYGIGMHLREVAVISA 360
 QY 360 NIOEAGCKKIFGSLAFLPESFDGPASTAPLOEQLQVFTLEITTYLYISAMPDNL 419
 DB 361 NVQBFPGCKKIFGSLAFLPESFDGPASTAPLOEQLQVFTLEITTYLYISAMPDNL 420
 QY 420 PDLSTFONLQVIRKGIILNNGAVSLTLOGIGISWLGIRSLREISGLALIHNTHLCPVHT 479
 DB 421 PDLSTFONLQVIRKGIILNNGAVSLTLOGIGISWLGIRSLREISGLALIHNTHLCPVHT 480
 QY 480 VPMDDLPRPHOALHTNARPEDE- CVBGSLACHOLCAHGCMGPRPTCCVNCOSPLRQ 538
 DB 481 VPMDDLPRPHOALHTNARPEDE- CVBGSLACHOLCAHGCMGPRPTCCVNCOSPLRQ 540
 QY 539 ECVSECRVLOGLPREVYNAHCLCPHPECOPONGSVTCGPBAPDQVACAHYKDPFCVA 598
 DB 541 ECVSECRVLOGLPREVYNAHCLCPHPECOPONGSVTCGPBAPDQVACAHYKDPFCVA 600
 QY 599 RCPGSGVPLSLMPTWKPPDBERGACOPPCINCTHSCVDLDDYGCAPBAPDQVACAHYKDPFCVA 658

DB 601 RCPGSGVPLSLMPTWKPPDBERGACOPPCINCTHSCVDLDDYGCAPBAPDQVACAHYKDPFCVA 660
 QY 659 VGIILVVLGVVFGILIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOQMIKKT 718
 DB 661 VQVLLFLLVVVGLIRBRQKIRKTYMRRLQETELVEPLTPSGAMPNOQMIKKT 720
 QY 719 ELRKVKVLSGAFGTVYKGIWIPGENVKIPVALIYVLENTSPKANKELDAVYVAGV 778
 DB 721 ELRKVKVLSGAFGTVYKGIWIPGENVKIPVALIYVLENTSPKANKELDAVYVAGV 780
 QY 779 SRVYSRLIGLCTSTVOLVTPOLMPYGLLDHYRNRGLSGODLNMCMQIAKMSYLED 838
 DB 781 SRVYSRLIGLCTSTVOLVTPOLMPYGLLDHYRNRGLSGODLNMCMQIAKMSYLED 840
 QY 839 VRLVHRDLAARNVLYKSNHXYITDPGLARLDIDETEVHADGKVPKIMWALESLRRR 898
 DB 841 VRLVHRDLAARNVLYKSNHXYITDPGLARLDIDETEVHADGKVPKIMWALESLRRR 900
 QY 899 FTHOSDWSYGVTVWELMTFGAKPYDGI PARBIPDLKGERLPORPCTIDVYIMVVC 958
 DB 901 FTHOSDWSYGVTVWELMTFGAKPYDGI PARBIPDLKGERLPORPCTIDVYIMVVC 960
 QY 959 WMISECRPRRELVSERSMARDPORFVYI ONEDLGASPLDSTFFYSLEDDMDGLV 1018
 DB 961 WMISECRPRRELVSERSMARDPORFVYI ONEDLGASPLDSTFFYSLEDDMDGLV 1020
 QY 1019 DAEELVYPOGFCFDPAPAGAGVHHRSSSTRSGGDLTLGLPSEBEAPRSLAPS 1078
 DB 1021 DAEELVYPOGFCFDPAPAGAGVHHRSSSTRSGGDLTLGLPSEBEAPRSLAPS 1080
 QY 1079 EGAGSDVFDGLMGAKGLOSPLTHDPSPLORESDEPTVPLPSTDTGVAPLTCSPPE 1138
 DB 1081 EGAGSDVFDGLMGAKGLOSPLTHDPSPLORESDEPTVPLPSTDTGVAPLTCSPPE 1140
 QY 1139 YVNOEDVPOPPSPREGPLPARAPAGATLERPKTSPKNGVYKVPAGAVENPEVLT 1198
 DB 1141 YVNOEDVPOPPSPREGPLPARAPAGATLERPKTSPKNGVYKVPAGAVENPEVLT 1200
 QY 1199 PQGGAAPQHPHPAPAFSPADNLVYWDQPPRGAAPSTFKGTPTANPEVYLGLDVPV 1255
 DB 1201 PREGTASPHSPAPAFSPADNLVYWDQPPRGAAPSTFKGTPTANPEVYLGLDVPV 1257

RESULT 3
 ERB2_MESAU
 ID ERB2_MESAU STANDARD; PRT; 1254 AA.
 AC Q60553; 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 ID (p185erbB2) (NEU proto-oncogene) (C-erbB-2).
 GN ERB2 OR NEU.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 RN NCBI_TaxID:10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Nerve;
 RX MEDLINE=94193007; PubMed=7908275;
 RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,
 RA Yamazaki Y., Ishikawa T.;
 RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";
 RL Gene 140:251-255(1994).
 CC -1- FUNCTION: Essential component of a neurogulin-receptor complex,
 CC although neurogulin does not interact with it alone. GP30 is a
 CC potential ligand for this receptor. Not activated by EGF, TGF-
 CC alpha and amphiregulin (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: Heterodimer with each of the other ERB2 receptors

(Potential). Interacts with PKCABP (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine residues.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.embnet.org/licenses/embnet.html> or send an email to license@ebi.ac.uk).

CC EMBL: D16295; BAA03801.1; -
CC PIR: I48161; I48161.
CC HSP: P13362; 1FGK.
DR InterPro: IPR000494; EGF_RL_domain.
DR InterPro: IPR006211; Furin-Like.
DR InterPro: IPR006212; Furin_repeat.
DR InterPro: IPR007019; Prot_Kinase.
DR InterPro: IPR001245; Tyr_Kinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00757; Furin-Like; 1.
DR Pfam: PF00069; Kinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_Kinase; 1.
DR SMART: SM00261; FU; 4.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Proto-oncogene; Disease mutation.
FT SIGNAL 1 21
FT CHAIN 22 1254
FT DOMAIN 22 652
FT TRANSMEM 653 675
FT DOMAIN 676 1254
FT DOMAIN 158 368
FT DOMAIN 472 644
FT DOMAIN 720 987
FT NP_BIND 726 734
FT BINDING 753 753
FT ACT_SITE 845 845
FT DISULFID 195 204
FT DISULFID 199 212
FT DISULFID 236 244
FT DISULFID 240 252
FT DISULFID 255 264
FT DISULFID 268 295
FT DISULFID 299 311
FT DISULFID 315 331
FT DISULFID 334 338
FT DISULFID 511 520
FT DISULFID 515 528
FT DISULFID 531 540
FT DISULFID 544 560
FT DISULFID 563 576
FT DISULFID 567 584
FT DISULFID 587 596
FT DISULFID 600 623
FT DISULFID 626 634
FT DISULFID 630 642
FT MOD_RES 1139 1139
FT MOD_RES 1247 1247
FT CARBOHYD 68 68
FT CARBOHYD 125 125
FT CARBOHYD 187 187
FT CARBOHYD 259 259

FT CARBOHYD 530 530 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 658 658 V -> E (IN ONCOGENIC NEU).
FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).
SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2BE1 CRC64;
Query Match 87.9%; Score 5993.5; DB 1; Length 1254;
Best Local Similarity 87.7%; Pred. No. 2.4e-309;
Matches 1101; Conservative 57; Mismatches 96; Indels 1; Gaps 1;
1 METALCWGILLALLPFGASTVCTGTMDLRLPASPETHLDLRLHYOCQVQVGNL 60
1 MELAMCGWGLLALLSPGASGTCTGTMDLRLPASPETHLDLRLHYOCQVQVGNL 60
61 ELTYLPTVASLPLDIDIEVGVYVLIANOVQVLEQLRLVIRGTQLEFDNALAVDNG 120
61 ELTYLPTVASLPLDIDIEVGVYVLIANOVQVLEQLRLVIRGTQLEFDNALAVDNR 120
121 PLNNTPTVATGASPGELRELOLRSLTEILKGVLIQRNPOLCYODTILMKDIPKONOLA 180
121 DELDNVTATGRTPEGLRELOLRSLTEILKGVLIQRNPOLCYODTILMKDIPKONOLA 180
181 LTLIDTNSRACHPCSPCKGSRCKGSESDCQSLTRYVCAGCARKCPLEPDCHEOC 240
181 PVDIDTNSRACHPCSPCKGSRCKGSESDCQSLTRYVCAGCARKCPLEPDCHEOC 240
241 AAGCTGPHGSDCLAFHNSGICELHCPALVYNTDTFESMPNEGRYTFGASCYACP 300
241 AAGCTGPHGSDCLAFHNSGICELHCPALVYNTDTFESMPNEGRYTFGASCYACP 300
301 YNYLSTDVSGCTLYVCPLENOBVTADGTQRCCKSKPCARVCYGLMEHLREAVATSN 360
301 YNYLSTDVSGCTLYVCPLENOBVTADGTQRCCKSKPCARVCYGLMEHLREAVATSN 360
301 YNYLSTDVSGCTLYVCPLENOBVTADGTQRCCKSKPCARVCYGLMEHLREAVATSN 360
301 YNYLSTDVSGCTLYVCPLENOBVTADGTQRCCKSKPCARVCYGLMEHLREAVATSN 360
361 IOEPAGCKKIGSLAFLESFSDGDPASMTAPLQEPOLQVFEFLERTGLYISAWPDSLP 420
361 IOEPAGCKKIGSLAFLESFSDGDPASMTAPLQEPOLQVFEFLERTGLYISAWPDSLP 420
421 DLSVQNTLOVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSSGLALIHNTHLCEVHTY 480
421 DLSVQNTLOVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSRLSSGLALIHNTHLCEVHTY 480
421 DLSVQNTLOVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSRLSSGLALIHNTHLCEVHTY 480
421 DLSVQNTLOVIRGRILHNGAVSLTQGLISWLGRLSRLSRLSRLSSGLALIHNTHLCEVHTY 480
481 PMDOLFRPHQALHTANRPEDECVGEGIALCHQLCARGHCWPGPTQVNCQSGFLRGQC 540
481 PMDOLFRPHQALHTANRPEDECVGEGIALCHQLCARGHCWPGPTQVNCQSGFLRGQC 540
481 PMDOLFRPHQALHTANRPEDECVGEGIALCHQLCARGHCWPGPTQVNCQSGFLRGQC 540
481 PMDOLFRPHQALHTANRPEDECVGEGIALCHQLCARGHCWPGPTQVNCQSGFLRGQC 540
541 VEECVLOGLRREYNNAHCLPCHRECPONQSVTCFGEBAQCVAACAYKDPFCVANC 600
541 VEECVLOGLRREYNNAHCLPCHRECPONQSVTCFGEBAQCVAACAYKDPFCVANC 600
541 VEECVLOGLRREYNNAHCLPCHRECPONQSVTCFGEBAQCVAACAYKDPFCVANC 600
541 VEECVLOGLRREYNNAHCLPCHRECPONQSVTCFGEBAQCVAACAYKDPFCVANC 600
601 PSGVSPDSTYPMKPEDEGACOPCPINCHSCVDLDERGCPAEBORASPAISILTVAG 660
601 PSGVSPDSTYPMKPEDEGACOPCPINCHSCVDLDERGCPAEBORASPAISILTVAG 660
601 PSGVSPDSTYPMKPEDEGACOPCPINCHSCVDLDERGCPAEBORASPAISILTVAG 660
601 PSGVSPDSTYPMKPEDEGACOPCPINCHSCVDLDERGCPAEBORASPAISILTVAG 660
661 ILLVVLGVVGLILKROOKIKRKYMRRLQETELVEPLTPSGAMPNQAQRILKETEL 720
661 ILLVVLGVVGLILKROOKIKRKYMRRLQETELVEPLTPSGAMPNQAQRILKETEL 720
661 ILLVVLGVVGLILKROOKIKRKYMRRLQETELVEPLTPSGAMPNQAQRILKETEL 720
661 ILLVVLGVVGLILKROOKIKRKYMRRLQETELVEPLTPSGAMPNQAQRILKETEL 720
721 RKVKYLSGAGRTYKGIWIPDEGENVKI PVAIKVLRRENTSPRANKELIDEAYVAGVSD 780
721 RKVKYLSGAGRTYKGIWIPDEGENVKI PVAIKVLRRENTSPRANKELIDEAYVAGVSD 780
721 RKVKYLSGAGRTYKGIWIPDEGENVKI PVAIKVLRRENTSPRANKELIDEAYVAGVSD 780
721 RKVKYLSGAGRTYKGIWIPDEGENVKI PVAIKVLRRENTSPRANKELIDEAYVAGVSD 780
781 YVSRLLGICLSTVQVLTQVLMPEYGLLHVRENRRGLSODLNNCMQIAKGSYLEDVR 840
781 YVSRLLGICLSTVQVLTQVLMPEYGLLHVRENRRGLSODLNNCMQIAKGSYLEDVR 840
781 YVSRLLGICLSTVQVLTQVLMPEYGLLHVRENRRGLSODLNNCMQIAKGSYLEDVR 840
781 YVSRLLGICLSTVQVLTQVLMPEYGLLHVRENRRGLSODLNNCMQIAKGSYLEDVR 840
841 LVHRDLAARNVLYSPNWKITDPGLARLLDIDETSYHADGKVPILKNALBSILRRRT 900
841 LVHRDLAARNVLYSPNWKITDPGLARLLDIDETSYHADGKVPILKNALBSILRRRT 900
841 LVHRDLAARNVLYSPNWKITDPGLARLLDIDETSYHADGKVPILKNALBSILRRRT 900
841 LVHRDLAARNVLYSPNWKITDPGLARLLDIDETSYHADGKVPILKNALBSILRRRT 900
901 HOSDWSYGVTVVLMATFGAKPYDGI PAREIPDLLEKGRLEPQPICTIDVYIMYKCMW 960
901 HOSDWSYGVTVVLMATFGAKPYDGI PAREIPDLLEKGRLEPQPICTIDVYIMYKCMW 960
901 HOSDWSYGVTVVLMATFGAKPYDGI PAREIPDLLEKGRLEPQPICTIDVYIMYKCMW 960
901 HOSDWSYGVTVVLMATFGAKPYDGI PAREIPDLLEKGRLEPQPICTIDVYIMYKCMW 960

Db 901 HOSDWSYGVTVWMLMFGANPYDGIIPARELIPDLLEKGERLPQPICTIDVIMNKKCM 960
 Qy 961 IDSCRPFRRLVSEFSHMAADPQRFVYIIONEDLGPASPLDSTFRSLLEDMDGLDYDA 1020
 Db 961 IDSCRPFRRLVSEFSHMAADPQRFVYIIONEDLGPASPLDSTFRSLLEDMDGLDYDA 1020
 Qy 1021 EBYLVPOGFFCPDPAPGAGGVHHRHRSSTRSGGSDLTGLLEPSEEDAPRSLPASEG 1080
 Db 1021 EBYLVPOGFFCPDPAPGAGGVHHRHRSSTRSGGSDLTGLLEPSEEDAPRSLPASEG 1080
 Qy 1081 AGSDVFPDDLGMGAAGKQSLPPTHDPSPLOKRSSEPTVPLPSEPTGCVYAPLTCSPQPEYV 1140
 Db 1081 AGSDVFPDDLGMGAAGKQSLPPTHDPSPLOKRSSEPTVPLPSEPTGCVYAPLTCSPQPEYV 1140
 Qy 1141 NOPDVFPQPSPREGEPPLPAARPAAGATLSPKTLSPGKGVVQVPAFGGAVENPEYLTPO 1200
 Db 1141 NOPDVFPQPSPREGEPPLPAARPAAGATLSPKTLSPGKGVVQVPAFGGAVENPEYLTPO 1200
 Qy 1201 GGAAPQPHPPAPFSPAPDNLYYMDODPPERGAPSPSTFKGTPTAENPEYLTGLDVPV 1255
 Db 1201 GGAAPQPHPPAPFSPAPDNLYYMDODPPERGAPSPSTFKGTPTAENPEYLTGLDVPV 1255

RESULT 4

ID EGFR HUMAN STANDARD; PRT; 1210 AA.
 AC P00533; 000688; 000732; P06268; 014225; Q92795; Q9B2S2; Q9GZX1;
 AC Q9H2C3; Q9H3C9; Q9UMD7; Q9UMD8; Q9UMG5;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor
 protein-tyrosine kinase ErbB-1).
 OS EGFR OR ERBB1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=84219729; PubMed=6328312;
 RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,
 RA Lee J., Varden Y., Liberman T.A., Schlessinger J., Downward J.,
 RA Mayes B.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;
 RT "Human epidermal growth factor receptor cDNA sequence and aberrant
 RT expression of the amplified gene in A431 epidermoid carcinoma cells";
 RL Nature 309:418-425 (1984).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Placenta;
 RX MEDLINE=95382957; PubMed=7554368;
 RA Ilekta J.V., Stark B.C., Scoccia B.;
 RT "Possible role of variant RNA transcripts in the regulation of
 RT epidermal growth factor receptor expression in human placenta";
 RL Mol. Reprod. Dev. 41:149-156 (1995).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Placenta;
 RX MEDLINE=97078686; PubMed=8918811;
 RA Reller J.L., Mahle N.J.;
 RT "A 1.8 kb alternative transcript from the human epidermal growth
 RT factor receptor gene encodes a truncated form of the receptor";
 RL Nucleic Acids Res. 24:4050-4056 (1996).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Placenta;
 RX MEDLINE=97256547; PubMed=9103388;
 RA Ilekta J.V., Gariti J., Niederberger C., Scoccia B.;
 RT "Expression of a truncated epidermal growth factor receptor-like
 RT protein (TEGFR) in ovarian cancer";
 RL Gynecol. Oncol. 65:36-41 (1997).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).

RC TISSUE=Placenta;
 RX MEDLINE=21100872; PubMed=11161793;
 RA Reller J.L., Thredgill D.W., Eley G.D., Strunk K.B., Danielson A.J.,
 RA Schehl Sinclair C., Pearshall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,
 RA Mahle N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative EGFR transcripts encoding truncated receptor
 RT isoforms";
 RL Genomics 71:1-20 (2001).
 RN [6]
 RP SEQUENCE OF 575-687 FROM N.A.
 RA Reller J.L., Thredgill D.W., Danielson A.J., Schehl C.M.,
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
 RA Mahle N.J.;
 RT "Human and mouse alternative EGFR transcripts encoding only the
 RT extracellular domain of the receptor";
 RL Submitted (Feb-1999) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 713-924 FROM N.A.
 RX MEDLINE=84196372; PubMed=6326261;
 RA Lin C.R., Chen W.S., Knutger W., Stolarczyk L.S., Weber W.,
 RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;
 RT "Expression cloning of human EGF receptor complementary DNA: gene
 RT amplification and three related messenger RNA products in A431
 RT cells";
 RL Science 224:843-848 (1984).
 RN [8]
 RP SEQUENCE OF 150-962 FROM N.A.
 RX MEDLINE=84245835; PubMed=6330563;
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,
 RA Roe B.A., Merlino G.T., Pastan I.;
 RT "Human epidermal growth factor receptor cDNA is homologous to a
 RT variety of RNAs overproduced in A431 carcinoma cells";
 RL Nature 309:806-810 (1984).
 RN [9]
 RP SEQUENCE OF 1028-1210 FROM N.A.
 RX MEDLINE=85046483; PubMed=6093780;
 RA Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,
 RA O'Malley B.W.;
 RT "Isolation of an evolutionarily conserved epidermal growth factor
 RT receptor cDNA from human A431 carcinoma cells";
 RL Biochem. Biophys. Res. Commun. 124:125-132 (1984).
 RN [10]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=86217333; PubMed=3329716;
 RA Haley J.D., Whittle N., Bennett P., Kitchington D., Ullrich A.,
 RA Waterfield M.D.;
 RT "The human EGF receptor gene: structure of the 110 kb locus and
 RT identification of sequences regulating its transcription";
 RL Oncogene Res. 1:375-396 (1987).
 RN [11]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=91107677; PubMed=1988448;
 RA Haley J.D., Waterfield M.D.;
 RT "Contributory effects of de novo transcription and premature
 RT transcript termination in the regulation of human epidermal growth
 RT factor receptor proto-oncogene RNA synthesis";
 RL J. Biol. Chem. 266:1746-1753 (1991).
 RN [12]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=85270438; PubMed=2991899;
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;
 RT "Characterization and sequence of the promoter region of the human
 RT epidermal growth factor receptor gene";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924 (1985).
 RN [13]
 RP SEQUENCE OF 540.
 RA Kohda D.;
 RL Submitted (SEP-1997) to the SWISS-PROT data bank.
 RN [14]
 RP RECEPTOR ACTIVITY;
 RX MEDLINE=84191554; PubMed=6325948;

[illegible]

DB 897 VMSVTVTWELMTFGSKPYDGIPIASEIISILEKGERLPOPPICITDVTMIMVCMIMAD 956
 QY 965 CRPRRELIVSESRNARDPQRFVVIQ-NEDLGPASPLDSTFYSLLEDMDKDLVADSEY 1023
 DB 957 SRPKRELILERSKXARBPQRYLVYQGBERMLPPTBSNFRALMDEEDMDVDADSEY 1016
 QY 1024 LVPOGQFPCCPDPAAGCGVHHRRSSSTRGCGDLTLGLBSEBAPRSLPASEGAS 1083
 DB 1017 LIPOGQFP-----SSPSTSRTPLLSLSLATS 1042
 QY 1084 DVFDSDLLGMAKAGLSLPTHDPSPLQRYSEDPTVPLPSET--DGVVALTCSPOPEYVN 1141
 DB 1043 N-NTVACIDRNAGSCPIKSDSFLQRYSDPTALTEDSIDTFL-----PVPEYIN 1094
 QY 1142 QPDVAPQPPSPREGPLPAAPAGATL-----ERPRTLSPGKGVKVDVAFAGAVENPEY 1196
 DB 1095 Q-----SVP-KRPAQSVQNPVYHNPQPLNPASRPHQD--PSTANGNPEY 1138
 QY 1197 L-TPQGAAPQPHPPAPAFNDLYWDQ-----DP-----PERGAPSTKQ 1239
 DB 1139 LNTVQ-----PTCVNSTRFDSPAHWAQKSHQISLDNPDYQODFPFKAKPNGIFKQ 1189
 QY 1240 TPTANPEYL 1249
 DB 1190 S-TANAEYL 1198

RESULT 5
 EGFR_MOUSE STANDARD; PRT; 1210 AA.
 ID EGFR_MOUSE
 AC Q01279;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxId:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RA MEDLINE=93026370; PubMed=1408137;
 RA Avivri A., Skorecki K., Yaron A., Givol D.;
 RT "Promoter region of the murine fibroblast growth factor receptor 2
 (bek/KGFR) gene.";
 RL Oncogene 7:1957-1962(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; and CD-1; TISSUE=Decidua, and Liver;
 RA MEDLINE=93126380; PubMed=7678348;
 RA Paria B.C., Das S.K., Andrews G.K., Day S.K.;
 RT "Expression of the epidermal growth factor receptor gene is regulated
 in mouse blastocysts during delayed implantation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RA Hibbs M.L.;
 RL Submitted (Apr-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B6/C3; TISSUE=Liver;
 RA MEDLINE=94170986; PubMed=8125255;
 RA Jenkine N.A., Lee D.C.;
 RT "The mouse waved-2 phenotype results from a point mutation in the EGF
 RT receptor tyrosine kinase.";
 RL Genes Dev. 8:399-413(1994).
 RN [5]
 RP SEQUENCE OF 1-714 FROM N.A.
 RC TISSUE=Brain;

RX MEDLINE=91232866; PubMed=2030916;
 RA Avivri A., Iax I., Ullrich A., Schlessinger J., Givol D., Morse B.;
 RT "Comparison of EGF receptor sequences as a guide to study the ligand
 RT binding site.";
 RL Oncogene 6:673-676(1991).
 RN [6]
 RP SEQUENCE OF 969-1117 FROM N.A.
 RC STRAIN=C3H;
 RA Eisinger D., P., Serrero G.;
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
 AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
 VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: Binds RPTK (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA
 CC synthesis, and cell proliferation.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; U03425; AA157899.1; -;
 DR EMBL; X59658; CA442319.1; -;
 DR EMBL; U06864; AA53029.1; -;
 DR EMBL; Z12608; CA478249.1; -;
 DR PIR; A53183; A53183.
 DR HSSP; P11362; 1P6K.
 DR MGD; MGI:95294; Egfr.
 DR GO; GO:0030139; C:cytosolic vesicle; IDA.
 DR GO; GO:0005622; C:intracellular; IDA.
 DR InterPro; IPR000494; EGFR_L_domain.
 DR InterPro; IPR006211; Furin-like.
 DR InterPro; IPR006212; Furin repeat.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF00069; Kinase; 1.
 DR Pfam; PF01030; Recep_L_domain; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMO0261; FU; 5.
 DR SMART; SMO0219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
 KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 1210
 FT FT 25 1210
 FT DOMAIN 25 647
 FT TRANSMEM 648 670
 FT DOMAIN 671 1210
 FT REPEAT 75 300
 FT REPEAT 390 600
 FT DOMAIN 1028 1071
 FT DOMAIN 714 981
 FT NP_BIND 720 728
 FT BINDING 747 747
 FT ACT_SITE 839 839
 FT DISTFTD 194 199
 FT DISTFTD 207 207
 FT DISULFD 215 223
 FT BY SIMILARITY.

Query Match: 46.2%; Score 3146; DB 1; Length 1210;
 Best Local Similarity: 49.7%; Pred. No. 6, 4e-159; Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;
 11 LLLALLPQGA--STOVCTGTDMKRLPASPEHLDMRLHYOGCOVVOGNTLETPLTN 68
 14 LITLCAAGALBEKKVCGSTNRLTQLGFEDHFLSLQRMNNGEVVGNLEITTVQKN 73
 69 ASLSFLQDIQEVGYVLIANOVROVPLQRLRIVRGTOIFEDNYALAVLNDGDPINNTP 128
 74 YDLSFLTKIQEYAVGVLLALNTVERIPLENLQIRGNALYENTYALALISN----- 124
 129 VTGAPGAGREIQLASLTETIILKGVLIQRMPOLCYODTILMDI---PHKNQALATI 184
 125 -YGTNRGTRELPMRLQELILGAVRSPNNPILCNNDITQMRDITVONVMSNMSMDL--- 180
 185 DTRRSRACHPSPMCKGRSCWSESSSEDCSLRTVCAAGCA--RCKGPLPTDCCHECAAG 243
 181 -QSHSSCKCPSCNGSCWGGEGENCQTKITICAOQCSHRCKRSRSDCCNCAAG 239
 244 CTGPKHSDCLACIHFNSGICELHCPALVTYNTDFESMNPNEGRTYFGASCYVACPYN 303
 240 CTGPRSDCLVCOKFDKATCKDTCPMLMNYPTTYQMDVNPBGKYSFGATCVKCKPRY 299
 304 LSTVGSCTLVPRPLNQETABEGTQRCCKSPCARVCYGLGEMELREAVTANIOE 363
 300 VVTDHSGSVRACGPDYEV--EEGRIKCKCKGCPCKVCNGSIGSPKOTLSINATNIG 358
 364 FAGCKKIFGSLAPLPSFGDDPSANTAPLOPBOLOVETLEBITGYLTVISAMPDLS 423
 359 FKCTAISDHLHLPAFAKGDSTRTPLDPRLEBLIKVKEITGGLLQAMPDNTDHL 418
 424 VFQNTQVIRGRIILHNGAYSLTQGLGISWLGRLSRLSGLALIHNTHLGFVHTVWD 483

419 AENELIIRGRKQGFSLAVANGNTISLGRSLKEISDGVITISGNNTLCYANTINK 478
 484 QLFRRPHOALLHTANRPDECVSGSLACIOLCARGHCWGPPTQCVCNCSQFLRGCEVES 543
 479 KLFGPNQKTKIMNRARCKCAVHVHVCNPLCSSSGCNGPBRDCVSCQNVSRGRCEVK 538
 544 CRVLQGLPREYVNAHCLPCHPECPONGSVTCFGEADQCVAAHYKDPFCVARNCPSC 603
 539 CNLEGEPRFEVNESECICHPHECLPQAMNITCTGRGPONCICQAHYIDGPHCVKCPAG 598
 604 VKPDLSTVPIKFPREBEACQCPCTNCHSCVTDLDKGPAPQARSPLTISIAVVGILL 663
 599 IGENNNTL-VWKYVADANNVCHLCHANCCTYGCAGPGLQCEVWPSGPKLPSIATGVGILL 657
 664 VVLGVNVEGI-LIKRQOKIRKYTRRRLLQETELVEPLTPSGAMPNOAMRLKETELRK 722
 658 FIVV-VALGIGLPMRRRIIVKRTIRLLQRELEVEPLTPSGEARNQALRIKETEFK 716
 723 VKVLGSAFGTVYKGIWIPDEGNVKI PVAIKVLRNNTSPKANKEILDEAYVNAVGSPPY 782
 717 IKVLGSAFGTVYKGLWIPDEGNVKI PVAIKELREATSPKANKEILDEAYVNAVSVNPHV 776
 783 SHLIGICTSTVQVLTQMLPVCCLLDHNRNRRGRGSDOLLWCMQIAGMSYLEDVRLV 842
 777 CRLLGICTSTVQVLTQMLPVCCLLDVYREHNDIGSOLLWCVQIAGMNYLEDRLV 836
 843 HDDLARVNLVSPHNVKITPGLARLLDIDETEVHADGKVPKIMMALESILRRFTQ 902
 837 HDDLARVNLVKTIPHNVKITPGLAKLGAEBKEHAEBGKVPKIMMALESILHRYTHQ 896
 903 SVMSYGVTVWELMTFGAPYDGIPIARIPDLLEKGERLPPTCTIDVYIMVNCWMD 962
 897 SVMSYGVTVWELMTFGSKPYDGIPIASDISILEKGERLPPTCTIDVYIMVNCWMD 956
 963 SECRFRFELVSEFRMARDPQRYVIO-NELDGASPLDSTFYSLLEDMDGIDLVAE 1021
 957 ADSRKRFEELLIFESKMRADPQRYLVIOGDEMHLPSPIDSNFYRLAMDEDMEDVDVAD 1016
 1022 EYLVPOQGFPCPPAPGAGWVHRRSSSTSGGDDLTGLPESEEARPSPLASEGA 1081
 1017 EYLVPOQGF-----NSBST-----SRTPLLSSLSA 1042
 1082 GSDVEDGDLGMAKAGLSLPTHDSPLQRYSEDPVLPSET--DGVAAPLTCSPQEP 1139
 1043 TSN-----NSTVACINRNSCRVXEDAPLQRYSSDPTGAVTEIDNIDAF-----VPEY 1092
 1140 VNQPDVRRQPSREGRPLPARPAGATL-----ERPKTISPGKGVYKQVFAFGAVENP 1194
 1093 VNQ-----SVP-KRPAGSVONPVYHNPPLHPAPORDLHYQN--PHSNVAVGNP 1136
 1195 EYL-TPOGAAVQHPHPAPSPAPDNLYYMD-----DP-----PERGAPSTF 1237
 1137 EYLTAAQ-----PTCLSSGFNSPALMIOKSHQSLNDPDIQDDFPFKETKNGIF 1187
 1238 KGTPTAENPEYGLDVP 1254
 1188 KG-PTAENAEYLRVAP 1203

RESULT 6
 ERB4 HUMAN
 ID ERB4 HUMAN
 AC 015303;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)
 DE (p10erbB4) (Tyrosine kinase-type cell surface receptor HBB4).
 GN ERB4 OR HBB4
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

RN NCBI_TaxID=9606;
RP SEQUENCE FROM N.A. (ISOFORM JM-A).
RC TISSUE=Breast carcinoma;
RX MEDLINE=31189574; PubMed=838326;
RA Plozman G.D., Caloucou J.-M., Whitney G.S., Green J.M., Carlton G.W.,
RA Roy L., Neubauer M.G., Shoyab M.;
RA "A novel junctional membrane domain isoform of HER4/Erbb4. Isoform-specific
RT tissue distribution and differential processing in response to
RT phorbol ester.";
RL J. Biol. Chem. 272:26761-26768(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).
RC TISSUE=Petal brain;
RX MEDLINE=97476287; PubMed=9334463;
RA Elenius K., Cortas G., Paul S., Choi C.J., Rio C., Plozman G.D.,
RA Klagsbrun M.;
RA "The neuroligin receptor ErbB-4 interacts with PDZ-containing proteins
RT at neuronal synapses.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:3596-3601(2000).
CC -1- FUNCTION: Specifically binds and is activated by neuroligin, NRG-
CC -2. NRG-3, heparin-binding EGF-like growth factor, betacellulin and
CC TNF $\alpha$ . Interaction with these factors induces cell differentiation.
CC Not activated by EGF, TGF- $\alpha$ , and amphiregulin.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: Homodimer or heterodimer with each of the other ERBB
CC receptors (potential). Interacts with the PDZ domain of the
CC synaptrophin SNTB2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Comment=The 2 isoforms differ functionally in their response to
CC phorbol ester. Isoform JM-A is processed but not isoform JM-B.
CC So, they respectively represent cleavable and noncleavable
CC forms of the receptor. Both isoforms are expressed in the
CC cerebellum, but only the isoform JM-B is expressed in the
CC heart;
CC Name=JM-A;
CC IsoId=O15303-1; Sequence=Displayed;
CC Name=JM-B;
CC IsoId=O15303-2; Sequence=VSP_002895;
CC -1- TISSUE SPECIFICITY: Expressed at highest levels in brain, heart,
CC kidney, in addition to skeletal muscle, parathyroid, cerebellum,
CC placental, spleen, testis and breast. Lower levels in thymus,
CC lung, salivary gland, and pancreas.
CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine
CC residues.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL, L07868, AB559446.1; -.
DR PIR, A47253, A47253.
DR HSPR, P11362, 1PGR.
DR GeneW, HGNC:3432, ERBB4.
DR MIM, 600543; -.
DR GO, GO:0005887, C:integral to plasma membrane; TAS.
DR GO, GO:0008283, P:cell proliferation; TAS.
DR GO, GO:0007275, P:development; TAS.

```

	DR	GO: GO:0007048; P oncogenesis; TAS.
	DR	InterPro: IPR000494; EGFR_L_domain.
	DR	InterPro: IPR006211; Furin_like.
	DR	InterPro: IPR006212; Furin_repeat.
	DR	InterPro: IPR000719; Prot_kinase.
	DR	InterPro: IPR001245; Tyr_kinase.
	DR	InterPro: IPR004019; YLP motif.
	DR	Pfam: PF00757; Furin-like; 1.
	DR	Pfam: PF00659; kinase; 1.
	DR	Pfam: PF01050; Recep_L_domain; 2.
	DR	Pfam: PF02757; YLP; 2.
	DR	PRINTS: PR00109; TYRKINASE.
	DR	ProDom: PD000001; Prot_kinase; 1.
	DR	SMART: SMO0261; Fu_5.
	DR	SMART: SMO0219; Tyrc; 1.
	DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
	DR	PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
	DR	PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
	KW	Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
	KW	Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
	KM	Alternative splicing.
FT	FT SIGNAL	1 25 POTENTIAL. RECEPTOR PROTEIN-TYROSINE KINASE ERBB-4. EXTRACELLULAR (POTENTIAL). CYTOPLASMIC (POTENTIAL).
FT	FT CHAIN	26 1308 26 651 CYS-RICH. CYS-RICH. CYTOSOL PLASMA MEMBRANE (POTENTIAL).
FT	FT TRANSMEM	652 675 676 1308 CYTOSOL PLASMA MEMBRANE (POTENTIAL).
FT	FT DOMAIN	186 334 496 633 PROTEIN KINASE. ATP (BY SIMILARITY). ATP (BY SIMILARITY).
FT	FT DOMAIN	718 985 724 732 NP BINDING. NP BINDING. ACT SITE.
FT	FT BINDING	751 751 843 843 BY SIMILARITY.
FT	FT ACT_SITE	843 843 BY SIMILARITY.
FT	FT DISULFID	189 197 BY SIMILARITY.
FT	FT DISULFID	193 205 BY SIMILARITY.
FT	FT DISULFID	213 221 BY SIMILARITY.
FT	FT DISULFID	217 229 BY SIMILARITY.
FT	FT DISULFID	230 238 BY SIMILARITY.
FT	FT DISULFID	234 246 BY SIMILARITY.
FT	FT DISULFID	249 258 BY SIMILARITY.
FT	FT DISULFID	262 289 BY SIMILARITY.
FT	FT DISULFID	293 304 BY SIMILARITY.
FT	FT DISULFID	308 323 BY SIMILARITY.
FT	FT DISULFID	326 330 BY SIMILARITY.
FT	FT DISULFID	503 512 BY SIMILARITY.
FT	FT DISULFID	507 520 BY SIMILARITY.
FT	FT DISULFID	523 532 BY SIMILARITY.
FT	FT DISULFID	536 552 BY SIMILARITY.
FT	FT DISULFID	555 569 BY SIMILARITY.
FT	FT DISULFID	559 577 BY SIMILARITY.
FT	FT DISULFID	580 589 BY SIMILARITY.
FT	FT DISULFID	593 614 BY SIMILARITY.
FT	FT DISULFID	617 625 BY SIMILARITY.
FT	FT DISULFID	621 633 BY SIMILARITY.
FT	FT MOD_RES	1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	FT MOD_RES	1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	FT MOD_RES	1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	FT MOD_RES	1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	FT CARBOHYD	138 138 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	174 174 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	181 181 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	253 253 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	358 358 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	410 410 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	473 473 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	495 495 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	548 548 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	576 576 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT CARBOHYD	620 620 N-LINKED (GLCNAC...) (POTENTIAL).
FT	FT VARSPLIC	626 648 IN ISOFORM JM-B.
SEQUENCE	1308 AA; 146807 MM; 55AABE09B5DB8761 CRC64; /FTId=VSP_002895.	

Query Match 44.1%; Score 3004.5; DB 1; Length 1308;
 Beet Local Similarity 45.6%; Pred. No. 2.1e-151;
 Matches 614; Conservative 183; Mismatches 375; Indels 175; Gaps 27;

9 WGLLALALPBGAA-----STVCGCTGDMKRLRPASPEPTHLDMLRHLVQGGVQVQGNLELY 64
 8 WYWSLLVAAAGTVQPSDSQSCAGTENKLSLSLDEQYRALUKTYENCEVWVGNLEITS 67
 65 LPTNALSPLFODIOEVQGYVLIAHQVROVPLQRLIVRGTOLEFBNVALAVLNDGPLAN 124
 68 IEHNDSLPLRSVRETVGYVALNQFRLPLENLAIIRGTLYERBALAIFLNRKXG 127
 125 NTPPYTASPGGLRELIQLSLTEILKGVLIQNPOLCTQDTILMKOIFHKNNQALTLI 184
 128 NF-----GLOBEGLKRLTEILNGVYVDQNKELCYADTITHMODIVRNWPSNLTLY 178
 165 DPNRSAPCHPCSPMCKSGRSMGSSSDCSLPTTVAGGC-ARCKPRLPTDCCBCCAG 243
 179 STNGSSGCRCHKCTG-RCKGFTENHCOTLRITVCAEQDGRCTGYVSDCCNRCAAG 237
 244 CTGPKHSDCLACLFPHNSGICELHCPALVTYNTDFESMNPREGRYTFGASCVTACPYNY 303
 238 CSGPDTDCFACMNFDSGACVTCQPTVYVNPPTQLHNHFAKATTYGAFCKYCKPHNF 297
 304 LSTDVSGCTLVCPRLHNOEVTABEDTQRCCKSKPCARVCYGLAMEHLREVAVTSANOE 363
 298 V-VDSSSCRACPSRMEV-EENGIMCKPCPTDIPCACDGICTGSLMGAQVDSNIDK 355
 364 FAGCKKIPCSLAFLEPSFDGDPASNPAPLOPBOLOVFELEETGLVLSAMDSLPJLS 423
 356 FINTCKINNLPLVTGTHGDPNALIEALIDPELDANFRVREITGLANTQSMPPNNTDS 415
 424 VFQNLQVINGRILHNGAVSLTLQSLGISWLGRLSLNELSGLALIHNTLHCFVHTVMD 483
 416 VPSNLVTIGRVLVSGSLILKQGITSLQFQSLKISAGNIYIINDNSLCYHTINNT 475
 484 QLFRRPHQALLHTANRPEDECVGEGIALCHOLCARHGMGPRPQCVCNCSOLFQGCVEE 543
 476 TLFSTINQRIIVIRDNKAKENCTAEQVCMHLCSSDCCMGPGPQCCSCRFSGRICBS 535
 544 CAVLGLPREVYNAHRCLECHPCOP-ONGSVTCFEPBADQCAVACHYDPPRCVARGCS 602
 536 CNLYDEEFEEFEGSICVECDPQCEMEQGLTCHBPGRNCTKCSHFQDPRVCVKCD 595
 603 GVKPDLSTYPIWKFPEBEGACQPCPINTCHSCVLDLDDKC-----PAEORASPL 651
 596 GIGQANSF--IFRYADPDRECHCHENCTQGCNGPSHDCIYYPWGHSTLPQHAH---- 649
 652 TSIIISAVV--GILLVVLGVFGLIKRRQOKIRKYMRLLOETELVEBLTPSGAMPNQ 709
 650 TELIAGVIGGLFIVLIVGLTAVVYRRRSIK-KQALRRFL-ETELVEBLTPSGAPNQ 707
 710 AQMRILKETELRKVKTLGSGAGCTYKGIWIPDGENVKIPVAIKVLRNTSPKANEILD 769
 708 AOLRIKETELRKVKTLGSGAGCTYKGIWIPDGENVKIPVAIKVLRNTSPKANEIFND 767
 770 BAYVMAVGSPPYVSRLLGICLTSTVOLVTLQMPYGCILDHVRNRRGLSGODLNNCMQI 829
 766 EALINASMHPHLVRLIGVCLSTIDVTLQMPHGCILBVEHAKNIGSOLLNNCVOI 827
 830 AKGMSYLEDVRLVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETEYHADGGKVPDIKM 889
 828 AKGMVLEERRRLVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETEYHADGGKVPDIKM 887
 890 ALESLIRREFTHOSDWSYGVYTWELMTTRAKAYDGIIPAREIDLLKGRRLPQPICTI 949
 888 ALECHYRKFTHOSDWSYGVYTWELMTTRAKAYDGIIPAREIDLLKGRRLPQPICTI 947
 950 DVMIMVWCKMTIDSECRPRELVSFSPNARAPORFVYIIONED-LGAPASPLDSTYRSI 1008
 948 DVMIMVWCKMTIDSECRPRELVSFSPNARAPORFVYIIONED-LGAPASPLDSTYRSI 1007
 1009 LEBDDMDGLVDABEYLVPQGFPCPDPAAGAGVHHRRHSSTRSGGGDLTLGLSPSBE 1068

DB 1008 LEBDDMDGLVDABEYLVP-QATNIPP-----ITYSRARIDSNRS-----EIGHSPRA 1055
 QY 1069 EAPRS-----PLAP-SEGAGSDVFDGLGMAAGLQ 1099
 DB 1056 YTPMNGNCFVRDGGFAAEGSVVPRAPTSNIPAPVAGATAEIFDSCCGTILRPV 1115
 QY 1100 SLPTHDPSPLOYSADPVPPLPS-----ETDGYAPLTCSPQPEYVQDPVPOPPSP 1152
 DB 1116 AHHVEDSSQYRSADPFVFAERSPRGELDEGVTWTPMRDKPKQEYLVPE----- 1167
 QY 1153 REGPLPAPRACATLERPKTSLSPKNGVYKVPAREGAVENPEYLTPOGGAAPQHPPPA 1212
 DB 1168 -ENPVSRR-----KNGDLQ-----ALDNEPYHNASNG-----PPKA 1198
 QY 1213 -----FSAFDNLVYWDOPPERGA--PPST 1236
 DB 1199 EDEYVNEPLVYNTFANTLGKAEYLLKNNILSMPEKAKADNDPYNHSLPRSTLQHPDY 1258
 QY 1237 FKGPPT-----AENPEYL 1249
 DB 1259 LQESTKYFYKONGRIRPIAENPEYL 1285

RESULT 7
 ERB4 RAT
 ID ERB4 RAT STANDARD; PRT; 1308 AA.
 AC 062956; 0942N7;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).
 GN ERB4 OR TYRO-2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98221155; PubMed=9553078;
 RA Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,
 Marchionni M.A., Kelly R.A.;
 RT "Neuregulin promote survival and growth of cardiac myocytes;
 RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult
 RT ventricular myocytes.";
 RT J. Biol. Chem. 273:10261-10269(1998).
 RN (2)
 RP SEQUENCE OF 848-901 FROM N.A.
 RC TISSUE=Sciatic nerve;
 RX MEDLINE=9122560; PubMed=2025425;
 RA Lai C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially
 RT expressed in the vertebrate nervous system.";
 RT Neuron 6:691-704(1991).
 RN (3)
 RP SEQUENCE OF 1031-1198 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Spinal cord;
 RX MEDLINE=97184212; PubMed=9030624;
 RA Carroll S.L., Miller M.L., Frohner P.W., Kim S.S., Corbett J.A.;
 RT "Expression of neuregulins and their putative receptors, ErbB2 and
 RT ErbB3, is induced during Wallerian degeneration.";
 RT J. Neurosci. 17:1642-1659(1997).
 CC -1- FUNCTION: Specifically binds and is activated by neuregulins, NRG-
 CC 2, NRG-3, heparin-binding EGF-like growth factor, betacellulin and
 CC NTAK. Interaction with these factors induces cell differentiation.
 CC Not activated by EGF, TGF-A, and amphiregulin (by similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: Homodimer or heterodimer with each of the other ERB
 CC receptors. Interacts with the PDZ domain of the synrophin SNR32
 CC (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: Preferentially expressed in the developing
 CC nervous system. Exhibits distinct and highly regionalized patterns
 CC of expression in the adult brain, where it is mainly found in the
 CC reticular nucleus of the thalamus. Very low levels in kidney, and
 CC heart.
 CC -1- PTH: Ligand-binding increases phosphorylation on tyrosine
 CC residues (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

CC EMBL, AF041838; AAD08899.1; -.
 CC EMBL, U52531; AAC53051.1; -.
 CC PIR, P10184; P10184.
 CC HSSP, P11362; 1FGK.
 CC InterPro: IPR000494; EGF_R_L_domain.
 CC InterPro: IPR006211; Furin_Like.
 CC InterPro: IPR006212; Furin_Repeat.
 CC InterPro: IPR007119; Tyr_Kinase.
 CC InterPro: IPR001245; Tyr_Kinase.
 CC InterPro: IPR004019; YLP_motif.
 CC Pfam: PF00757; Furin_Like; 1.
 CC Pfam: PF00069; pkinase; 1.
 CC Pfam: PF01030; Recep_L_domain; 2.
 CC Pfam: PF01757; YLP; 2.
 CC Pfam: PF00001; Proct_Kinase; 1.
 CC SMART; SMO0261; FU; 5.
 CC SMART; SMO0219; Tyrc; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC Trnasembatene; Glycoprotein; Multigene family; Receptor; Signal;
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
 CC SIGNAL 1 25
 CC CHAIN 26 1308 RECEPTOR PROTEIN-TYROSINE KINASE EMBB-4.
 CC DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 652 675 POTENTIAL.
 CC DOMAIN 186 1308 CYTOPLASMIC (POTENTIAL).
 CC DOMAIN 496 633 CYS-RICH.
 CC DOMAIN 718 985 CYS-RICH.
 CC NP_BIND 724 732 ATP (BY SIMILARITY).
 CC BINDING 751 751 ATP (BY SIMILARITY).
 CC ACT_SITE 843 843 BY SIMILARITY.
 CC DISULFID 189 197 BY SIMILARITY.
 CC DISULFID 193 205 BY SIMILARITY.
 CC DISULFID 213 221 BY SIMILARITY.
 CC DISULFID 217 229 BY SIMILARITY.
 CC DISULFID 230 238 BY SIMILARITY.
 CC DISULFID 249 258 BY SIMILARITY.
 CC DISULFID 262 289 BY SIMILARITY.
 CC DISULFID 293 304 BY SIMILARITY.
 CC DISULFID 308 323 BY SIMILARITY.
 CC DISULFID 326 330 BY SIMILARITY.
 CC DISULFID 503 512 BY SIMILARITY.
 CC DISULFID 507 520 BY SIMILARITY.
 CC DISULFID 523 532 BY SIMILARITY.
 CC DISULFID 536 552 BY SIMILARITY.
 CC DISULFID 555 569 BY SIMILARITY.
 CC DISULFID 559 577 BY SIMILARITY.
 CC DISULFID 580 589 BY SIMILARITY.
 CC DISULFID 593 614 BY SIMILARITY.
 CC DISULFID 617 625 BY SIMILARITY.
 CC DISULFID 621 633 BY SIMILARITY.
 CC MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 CC MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 CC MOD_RES 1168 1168

FT MOD_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 548 548 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 1062 1062 S -> N (IN REF. 3).
 FT CONFLICT 1080 1082 PYT -> SYR (IN REF. 3).
 SQ SEQUENCE 1308 AA; 146957 MW; D944B096A08841 CRC64;

Query Match 43.9%; Score 2889; DB 1; Length 1308;
 Beet Local Similarity 46.2%; Pred. No. 1.4e-150;
 Matches 612; Conservative 194; Mismatches 402; Indels 118; Gaps 27;

1 MEIA-ALCRWGLL--ALLPGAASTVCTGDMRLRASPETHLDMLRHLYOGCQVQ 57
 1 MKLNGLVWWSGLVAARTVQPSASQVCACTENLSSLSDEQGYRALRKYENCEVVM 60
 58 GLELEITYLPTNASLSEFLDIOEVQVLIANQVQVPLQRLIRYGTQLFEDNYALAVL 117
 61 GLELEITSIEHNRDLSEFLRSIREVTGYVALNQFRLPLENRIIRGTLYEDRYALAI 120
 118 DNGDPLNNTTPTVQASPGGLARELOASLTETLKGVULQORPOLCYODITLKKDIFHKRN 177
 121 LNYRKGNF-----GLDELGLKNTLELNGVYVDQNFLEYADITLHODIRNPM 171
 178 QALTLIDITNRSRACHPCSPCKSGRCSSESEDCSLTRYACAGGC-ARCKGPLPTDCC 236
 172 PSNMTLVITIGSSGGRCHKCTG-RCMGPTENHCOTILTRYCACGDCRCGPVYSDCC 230
 237 HQCAAGTGTPRHSDCLALPHNHSIGCELPALVTYNTDTPESMPNDEGRYTFGASCV 296
 231 HRECAAGSGSDPTDPCFACNMFNDSGACVTCQPFVYVPTFQLEHNFNATYGAFCV 290
 237 TACPTNYLSTVGSTLYVCPHINQVTAEDQORBEKSKPCARCYGLGMEHREVRVAV 356
 291 KCPNPNFV-VSSSCVRACPSCKMEV-BENGIKMKPCPTDICKPKCDGIGTGLMSAQTV 348
 357 TSANIOEPAAGCKKIGSLAFLESPFGDPASNTAFLQPEQLQVFTLEITGLYISAMP 416
 349 DSNIDKINCKIKINGLILFLVTGIGHDPYNAIDAIPEKANVFRVRIITGFLNQTMP 408
 417 DSLPDLVYFONLQVIRGRILHNGAVSLTQGLGISMLGSLRLSGSLALIHNTLCP 476
 409 PMTDPVSFVSNLVTIGRVLVSGSLILTKQGGITSLQFOSLKEISAGNIYITDNNLCY 468
 477 VHTVWDOLFRRPHOALHTANRPREDCEYBGLACHOLCARHNCGRGTQCNCSQFLR 536
 469 YHTIMWTLTFSVNRIRYIDNRRAENCTABEMVNCNHLCSNDGCMGPGDQCLSCRFES 528
 537 GGECEVECRVLOGLEREYNARHCLPCHPECP-ONGSVTCGPEADQCVACAHYDPPF 595
 529 GKICIESCNLDGERREBENGSI CYECDSQCEKMDGLITGCPGPDNCTKSHFPDGN 568
 596 CYARCPGKPDLSYMPWKPFDEBGAQPCPINCTHSCVDLDDKGC-----PA 644
 589 CYEKCPDVLQGANSE--IFKYADQRECHPCPNCTOGCNGPFSHDIYYPTGSHTLQ 646
 645 EGRAPPLSIISAVV--GILVVVGVVPGIILKXROQKIRKYTRMRLLQETELVPLP 702
 647 HAR---TPLLAAGVIGGLFVIALTLFAVYVRRKSK-KRALRFL-ETELVPLP 700
 703 SGAMPNQAOMRIKTELARKVVLSSGAFGYKGIWIPDGNVKIPVAIKVLRNTSPK 762
 701 SGTANQAOMRIKTELARKVVLSSGAFGYKGIWVPEGTVKIPVAIKILNETGPK 760
 763 AKKEILDEAYVAGVSPVSRLLGICLTSTVOLVTQMPYGCGLDHDVRENKRLGSQDL 822

```

Db 761 ANVEPMEDALMAVDHPLVLLGLVCLSPITQVLTQVLMHOCLEVEYHEKDNIGSOL 820
Cc 823 LNMCHQKNGKSYLDVRLVHDLAARNVLYKSPHNVKTTDGLARLIDIDETEXHADGG 882
Cc 821 LNMCCQKNGKMYLBERLVHDLAARNVLYKSPHNVKTTDGLARLIDIDETEXHADGG 880
Cc 883 KVPKMMALLESILRRRFTQSDVMSYGVTVWELMTFGAKPYDQIPAREIPDLLKEGERLP 942
Cc 881 KMPKMMALLESILRRRFTQSDVMSYGVTVWELMTFGAKPYDQIPAREIPDLLKEGERLP 940
Cc 943 OPICTIVYMMVMKCMWIDSECRPRFRLVSEFPMARDPQRFVYVIONED-LGPAASPLD 1001
Cc 941 OPICTIVYMMVMKCMWIDSECRPRFRLVSEFPMARDPQRFVYVIONED-LGPAASPLD 1000
Cc 1002 STFYSLLEDDMDGLVDAEELVLPQGGFCFCDP-----ARG 1038
Cc 1001 SKFFQNLDEEDLEMDDEELVLP-QAFNIPPIYTSRTRIDSNRSEIGHSPPRAYTPM 1059
Cc 1039 AGGVVHHRHRSSTRSGGGDLTLGLEPSEBEAPRSPLASEGSDVDPDGLMGAKGL 1098
Cc 1060 SGSGVYVDDGSGATOG--MEMPTATATSTIPBAVVA--QGATLAMPDSCNGTLRKP 1114
Cc 1099 GSLPTHDSPLQRYSDPVPPLPS-----ETDGVAPLTCSPQPEYVNOQDVAPQPPS 1151
Cc 1115 VVPHVQSDSSTQRYSDPVPPLPS-----AFGAVENPEYLTP 1199
Cc 1152 PREGPLPARPAGATLERPK---TLSPGKNGVVOVF-----AFGAVENPEYLTP 1199
Cc 1175 RONGLOA-----LDNPEYHSSASGPK---AEDEYVNEPLYLNTFNALGNMEYMKM 1224
Cc 1200 OGGAPOHPPEPAPFAPNLYWDDPPEPGA--PPSTFKGTP-----A 1243
Cc 1225 SLASVPE-----KAKKAFDNPDMYHNSLPRSTGLQHPDYLQGYSTKYFYKNGRIRPIYA 1279
Cc 1244 ENPEYL 1249
Cc 1280 ENPEYL 1285

```

RESULT 8

```

XMRK_X1PMA STANDARD; PRT; 1167 AA.
ID XMRK_X1PMA
AC P13388;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Melanoma receptor protein-tyrosine kinase precursor (EC 2.7.1.112).
GN XMRK OR TV.
OS Xiphophorus maculatus (Southern platyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Perciformes; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OC NCBI_TaxID=8083;
OX 11
RN SEQUENCE FROM N.A.
RP MEDLINE=90015140; PubMed=2797166;
RX Wilbrode J., Adam D., Maltzschek B., Maueier W., Raulf F.,
RA Telling A., Robertson S.M., Schartl M.;
RT "Novel putative receptor tyrosine kinase encoded by the melanoma-
RT inducing Tu locus in Xiphophorus.";
RL Nature 341:415-421 (1998).
RP REVISION TO 515.
RA Schartl M.;
RL Submitted (Jun-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DISEASE: INVOLVED IN PIGMENT CELLS MALIGNANT MELANOMAS.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

```

```

Cc -----
Cc This SWISS-Prot entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (see http://www.isb-sib.ch/announce/
Cc or send an email to license@sib-bib.ch).
Cc -----
Cc DR EMBL; X16891; CAA34770.2; -
Cc DR PIR; S06142; S06142.
Cc DR HSP; P1362; 1REK.
Cc DR InterPro; IPR00494; EGFR_L_domain.
Cc DR InterPro; IPR006211; Furin-like.
Cc DR InterPro; IPR006212; Furin repeat.
Cc DR InterPro; IPR00719; Prot_kinase.
Cc DR InterPro; IPR01245; Tyr_kinase.
Cc DR Pfam; PF00757; Furin-like; 1.
Cc DR Pfam; PF0069; Kinase; 1.
Cc DR Pfam; PF01030; Recep_L_domain; 2.
Cc DR PRINTS; PS0109; TYRKINASE.
Cc DR ProDom; PD000001; Prot_kinase; 1.
Cc DR SMART; SM00261; Fu; 5.
Cc DR SMART; SM00219; Tyrc; 1.
Cc DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
Cc DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
Cc DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
Cc KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
Cc KM Tyrosine-protein kinase; ATP-binding; phosphorylation; Proto-oncogene.
Cc FT SIGNAL 1 25
Cc FT CHAIN 26 1167
Cc FT 26 1167 MELANOMA RECEPTOR PROTEIN-TYROSINE
Cc FT 26 1167 KINASE.
Cc FT 26 1167 EXTRACELLULAR (POTENTIAL).
Cc FT 642 642 POTENTIAL.
Cc FT 643 665 CYTOPLASMIC (POTENTIAL).
Cc FT 666 1167 PROTEIN KINASE.
Cc FT 710 977 ATP (BY SIMILARITY).
Cc FT NP_BIND 716 724 ATP (BY SIMILARITY).
Cc FT BINDING 743 743
Cc FT ACT_SITE 835 835
Cc FT DISULFID 195 204
Cc FT DISULFID 199 212
Cc FT DISULFID 220 228
Cc FT DISULFID 224 236
Cc FT DISULFID 237 245
Cc FT DISULFID 241 253
Cc FT DISULFID 256 265
Cc FT DISULFID 269 296
Cc FT DISULFID 300 311
Cc FT DISULFID 315 330
Cc FT DISULFID 333 337
Cc FT DISULFID 504 513
Cc FT DISULFID 508 521
Cc FT DISULFID 524 533
Cc FT DISULFID 537 553
Cc FT DISULFID 556 569
Cc FT DISULFID 560 577
Cc FT DISULFID 593 615
Cc FT DISULFID 618 626
Cc FT DISULFID 622 634
Cc FT CARBOHYD 114 114
Cc FT CARBOHYD 144 144
Cc FT CARBOHYD 201 201
Cc FT CARBOHYD 356 356
Cc FT CARBOHYD 365 365
Cc FT CARBOHYD 398 398
Cc FT CARBOHYD 417 417
Cc FT CARBOHYD 501 501
Cc FT CARBOHYD 576 576
Cc FT CARBOHYD 621 621
Cc SO SEQUENCE 1167 AA; 129934 MW; 4793B0749DCD55A CRC64;

```

Query Match 40.0%; Score 2724.5; DB 1; Length 1167;
Best Local Similarity 45.5%; Pred. No. 1.1e-136;

Matches 577; Conservative 164; Mismatches 389; Indels 137; Gaps 27;

4 AALCRGGLLALLPFGAASST---QVCTGTDMLRLPAPETHTLDMRLHLYOCQOVQGN 59
 8 AALLQ--LLLVLSISRCSTDPDRKVCQOSTSNQMT---LDNHYLKMKNVSGCNVLEN 62
 60 LELTYLPLNVAISPLDIOEVQGYVLIANOVROVPLQRLIVRGTFQEDVYALALDN 119
 63 LEITYQEQDLSPLQIEVGGYVLIANNEVSTIPLVNLRLIRGNLYEGNFTLLVMSN 122
 120 GDPNNTTPVTCASFGCLRELQRLSTELLKGVLIQRPOLCYODTILMKOIFHKNNL 179
 123 YOK-PPSSP--DYOVGKQLQSLNLTLLSGVKSHPNPLCNVETIMWMDI VDKTSP 179
 180 ALLTIDTRNSRACHPSPWCKGSRGWSGSSSDQSLTRTVCAAGC--ARCKGFLPTDCHE 238
 180 TMLLPLHAFBRQCKQCDHGCNAGSCAPRGHCQKTKLLCAEQCNRRCRGKRPIDCCNB 239
 239 OCAAGCTGPKHSDCLACLPHNSGICELACPALVTYNTDTFESMPNREGRYTFGASCYA 298
 240 HCAGGCTGPRATDCLACRDFNDGCTKOTCPPKTYDIYSHQVNDPNIKYTFGAACVKE 299
 299 CRYNTLSTDVGSCTVCPLEHNOETABDGTQRCCKSKCAVCYGLGMEHLREVAATVS 358
 300 CFSNTVYVE--GACVSSCSAGMLEVD--ENGRCKCPDGVCPKVCDDIGISLNTIAVNS 357
 359 ANIOBPAGCKKIFGSLAFPESSPDGDPANNTAPLOPELOVFETLEITGYLYISAMPDS 418
 358 TNIRSFNSTKINGDILNRSFBDPHYKIGTMDBEHLMLTVEITGYLYIVMMFEN 417
 419 LPLDSVFOQLQVIRGRIILNNGAYS--LTLQGLSISWGLSRLEBSGLALIHNTHLCEV 477
 418 MTSLSVFOQLLEIRGRTTSRGRSFVVQVVRHLQMLGLASLKEVSGANVILKNLQLRYA 477
 478 KIVPMDOLFRMHQALHTAARPEDECVBGLAQHLCARGHCWGGPFOCNCQSPFLAG 537
 478 NTINNRRLPRSEDQSI EYDART-----ENOTCNNECSBDQWGMGPWCVSCILVNDG 530
 538 QECVEBCRLQGLPREYVNAARHCLPCHPECOPONGSVTFEGPEADQCVACAHYKDPFCV 597
 531 GRCVASCNLQGBPREAOVDGRVCQCHQGLVOTDSLTTCYGEPAANCSSAHPQDPQCI 590
 598 AACPSGVKEDLSYMPIWKPDEBGAQCPPCINCTHSCVDLDDKCAPAORASPLTSIISA 657
 591 PCPFGILGDDPTL-IMKYADKMGQCPHQNCTQCSGPGISGCGD--IVSHSLAVGL 648
 658 VVGLILVVLGVFGLIKRQOKIRKYMRLQSTELVEPLTPSGAMPNQAKRIKE 717
 649 VSGLLITVIVALLIVLLRRRRIRK-RKRTIRCLQEKELVEPLTPSGOAPNOAFLRIKE 707
 718 TELRKVKVGSAGFGVYVYKIMIPDGENYKIPAIIVLRNTPSPKANKELDEAYYMACV 777
 708 TEPKDQVLSGAGFGVYVYKIMIPDGENYKIPAIIVLRNTPSPKANKELDEAYYMACV 767
 778 GSPYVSRLLGICLSTVQVLTQMLPQGLLDHVENRGLSGQDLNMCQIAKMSYLE 837
 768 DHPHCRLIGLICLTSAVQVLTQMLPQGLLDHVENRGLSGQDLNMCQIAKMSYLE 827
 838 DVLVLRDLAARVYVKSPPHVKITDPGLAKLLIDITETVHADGKVPYIKMALESILKR 897
 828 EHLVLRDLAARVYVKSPPHVKITDPGLAKLLIDITETVHADGKVPYIKMALESILKR 887
 898 RETHOSDVSYGVTEMLMTFGAKPYDGI-PAREIPULEKGERLPPPICTIDVYIMYK 957
 888 TYTHOSDVSYGVTEMLMTFGAKPYDGI-PAREIPULEKGERLPPPICTIDVYIMYK 947
 958 CMVIDSECPRRRELVSFSSRMARDQRFVVIQNEBLAGPASPIDSTFYSSLLEDDMDGL 1017
 948 CMVIDSSRRPRRELVSFSSRMARDQRFVVIQNEBLAGPASPIDSTFYSSLLEDDMDGL 1002
 1018 VDAEYLIVQGGFCFCDPAPAGAGCAHHHRSSSTSSGGGDLTLGLEPSEBEPAPPLAP 1077
 1003 VDAEYLIVQGGFCFCDPAPAGAGCAHHHRSSSTSSGGGDLTLGLEPSEBEPAPPLAP 1025

1078 SEGASDVFDGDLGMAKGLQSLPETHDPSPLQRYSEDPV-PLPSETDGVVAPLTCSPO 1136
 1026 PFGH-----FVRENSTLRNISPPTQVLALEKDLGDH----- 1056
 1137 PEYVQPPVPRPOP-----PSPRE-----GFLP--AARPAQTLEPRKTLSSGKGVKD 1183
 1057 -EYVQPPSETSSRLSDIYNPYEDLTGQWGVSSLSQEAETNFSRPEYLTNTNNSL--- 1112
 1184 VFAPGAVENPEYLTPOGGAAPQHPAPFAPFONLYYWDODPPERGAPSTFKCTPTA 1243
 1113 PLVSSGSMDDPY---QAG-----YQAAF-----LPQTGALTNGMFLPPAA 1150
 1244 ENPEYLG 1250
 1151 ENPEYLG 1157

RESULT 9
 ERB3 HUMAN
 ID ERB3 HUMAN . STANDARD; PRT; 1342 AA.
 AC P21860;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)
 DE (c-erbB3) (Tyrosine kinase-type cell surface receptor HER3).
 GN ERB3 OR HER3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=90083234; PubMed=2687875;
 RA Kraus M.H., Issing W., Mixt T., Popescu N.C., Aaronson S.A.;
 RT "Isolation and characterization of ERB3, a third member of the
 RT ERBB/epidermal growth factor receptor family: evidence for
 RT overexpression in a subset of human mammary tumors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:9193-9197(1989).
 RN (2)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=90311312; PubMed=2164210;
 RA Plozman G.D., Whitney G.S., Neubauer M.G., Green J.M., McDonald V.L.,
 RA Todaro G.J., Shoyab M.;
 RT "Molecular cloning and expression of an additional epidermal growth
 RT factor receptor-related gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4905-4909(1990).
 RN (3)
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=placenta;
 RX MEDLINE=93282822; PubMed=7685162;
 RA Katch M., Yazaki Y., Sugimura T., Terada M.;
 RT "c-erbB3 gene encodes secreted as well as transmembrane receptor
 RT tyrosine kinase.";
 RL Biochem. Biophys. Res. Commun. 192:1189-1197(1993).
 CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAK.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
 CC (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG FORM) AND
 CC SECRETED (SHORT FORM).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1 Synonym=long form; Sequence=displayed;
 CC Name=2 Synonym=short form; Sequence=displayed;
 CC Name=2 Synonym=short form; Sequence=VSP 002893; VSP 002894;
 CC IsoId=P21860-2; Sequence=VSP 002893; VSP 002894;
 CC -1- TISSUE SPECIFICITY: EPITHELIAL TISSUES AND BRAIN.
 CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
 CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES

CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
 CC PHOSPHATIDYLINOSITOL 3-KINASE (BY SIMILARITY).
 CC -1- DISEASE: OVEREXPRESSED IN A SUBSET OF HUMAN MAMMARY TUMORS.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; M29366; AAA5790.1; -;
 CC EMBL; M34309; AAA5791.1; -;
 CC EMBL; S61953; AAA26935.1; -;
 CC PIR; A36223; A36223.
 CC PIR; JH0803; JH0803.
 CC PDB; 1M6B; 23-AUG-02.
 CC GeneW; HGNC:3431; ERBB3.
 CC MIM; 190151; -;
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0005906; E:epidermal growth factor receptor activity; TAS.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC InterPro; IPR000494; EGFR_L_domain.
 CC InterPro; IPR006212; Furin-Like.
 CC InterPro; IPR000719; Prot_Kinase.
 CC InterPro; IPR001245; Tyr_Kinase.
 CC Pfam; PF00757; Furin-like; 1.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF01030; Recep_L_domain; 2.
 CC PRINTS; PRO0109; TYRKINASE.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00261; FU_5.
 CC SMART; SM00219; TYKc; 1.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 CC PROSITE; PS00109; PROTEIN KINASE TYR; FALSE_NEG.
 CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 CC Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 CC Alternative splicing; 3D-structure.
 CC K1
 CC K2
 CC K3
 CC K4
 CC K5
 CC K6
 CC K7
 CC K8
 CC K9
 CC K10
 CC K11
 CC K12
 CC K13
 CC K14
 CC K15
 CC K16
 CC K17
 CC K18
 CC K19
 CC K20
 CC K21
 CC K22
 CC K23
 CC K24
 CC K25
 CC K26
 CC K27
 CC K28
 CC K29
 CC K30
 CC K31
 CC K32
 CC K33
 CC K34
 CC K35
 CC K36
 CC K37
 CC K38
 CC K39
 CC K40
 CC K41
 CC K42
 CC K43
 CC K44
 CC K45
 CC K46
 CC K47
 CC K48
 CC K49
 CC K50
 CC K51
 CC K52
 CC K53
 CC K54
 CC K55
 CC K56
 CC K57
 CC K58
 CC K59
 CC K60
 CC K61
 CC K62
 CC K63
 CC K64
 CC K65
 CC K66
 CC K67
 CC K68
 CC K69
 CC K70
 CC K71
 CC K72
 CC K73
 CC K74
 CC K75
 CC K76
 CC K77
 CC K78
 CC K79
 CC K80
 CC K81
 CC K82
 CC K83
 CC K84
 CC K85
 CC K86
 CC K87
 CC K88
 CC K89
 CC K90
 CC K91
 CC K92
 CC K93
 CC K94
 CC K95
 CC K96
 CC K97
 CC K98
 CC K99
 CC K100
 CC K101
 CC K102
 CC K103
 CC K104
 CC K105
 CC K106
 CC K107
 CC K108
 CC K109
 CC K110
 CC K111
 CC K112
 CC K113
 CC K114
 CC K115
 CC K116
 CC K117
 CC K118
 CC K119
 CC K120
 CC K121
 CC K122
 CC K123
 CC K124
 CC K125
 CC K126
 CC K127
 CC K128
 CC K129
 CC K130
 CC K131
 CC K132
 CC K133
 CC K134
 CC K135
 CC K136
 CC K137
 CC K138
 CC K139
 CC K140
 CC K141
 CC K142
 CC K143
 CC K144
 CC K145
 CC K146
 CC K147
 CC K148
 CC K149
 CC K150
 CC K151
 CC K152
 CC K153
 CC K154
 CC K155
 CC K156
 CC K157
 CC K158
 CC K159
 CC K160
 CC K161
 CC K162
 CC K163
 CC K164
 CC K165
 CC K166
 CC K167
 CC K168
 CC K169
 CC K170
 CC K171
 CC K172
 CC K173
 CC K174
 CC K175
 CC K176
 CC K177
 CC K178
 CC K179
 CC K180
 CC K181
 CC K182
 CC K183
 CC K184
 CC K185
 CC K186
 CC K187
 CC K188
 CC K189
 CC K190
 CC K191
 CC K192
 CC K193
 CC K194
 CC K195
 CC K196
 CC K197
 CC K198
 CC K199
 CC K200
 CC K201
 CC K202
 CC K203
 CC K204
 CC K205
 CC K206
 CC K207
 CC K208
 CC K209
 CC K210
 CC K211
 CC K212
 CC K213
 CC K214
 CC K215
 CC K216
 CC K217
 CC K218
 CC K219
 CC K220
 CC K221
 CC K222
 CC K223
 CC K224
 CC K225
 CC K226
 CC K227
 CC K228
 CC K229
 CC K230
 CC K231
 CC K232
 CC K233
 CC K234
 CC K235
 CC K236
 CC K237
 CC K238
 CC K239
 CC K240
 CC K241
 CC K242
 CC K243
 CC K244
 CC K245
 CC K246
 CC K247
 CC K248
 CC K249
 CC K250
 CC K251
 CC K252
 CC K253
 CC K254
 CC K255
 CC K256
 CC K257
 CC K258
 CC K259
 CC K260
 CC K261
 CC K262
 CC K263
 CC K264
 CC K265
 CC K266
 CC K267
 CC K268
 CC K269
 CC K270
 CC K271
 CC K272
 CC K273
 CC K274
 CC K275
 CC K276
 CC K277
 CC K278
 CC K279
 CC K280
 CC K281
 CC K282
 CC K283
 CC K284
 CC K285
 CC K286
 CC K287
 CC K288
 CC K289
 CC K290
 CC K291
 CC K292
 CC K293
 CC K294
 CC K295
 CC K296
 CC K297
 CC K298
 CC K299
 CC K300
 CC K301
 CC K302
 CC K303
 CC K304
 CC K305
 CC K306
 CC K307
 CC K308
 CC K309
 CC K310
 CC K311
 CC K312
 CC K313
 CC K314
 CC K315
 CC K316
 CC K317
 CC K318
 CC K319
 CC K320
 CC K321
 CC K322
 CC K323
 CC K324
 CC K325
 CC K326
 CC K327
 CC K328
 CC K329
 CC K330
 CC K331
 CC K332
 CC K333
 CC K334
 CC K335
 CC K336
 CC K337
 CC K338
 CC K339
 CC K340
 CC K341
 CC K342
 CC K343
 CC K344
 CC K345
 CC K346
 CC K347
 CC K348
 CC K349
 CC K350
 CC K351
 CC K352
 CC K353
 CC K354
 CC K355
 CC K356
 CC K357
 CC K358
 CC K359
 CC K360
 CC K361
 CC K362
 CC K363
 CC K364
 CC K365
 CC K366
 CC K367
 CC K368
 CC K369
 CC K370
 CC K371
 CC K372
 CC K373
 CC K374
 CC K375
 CC K376
 CC K377
 CC K378
 CC K379
 CC K380
 CC K381
 CC K382
 CC K383
 CC K384
 CC K385
 CC K386
 CC K387
 CC K388
 CC K389
 CC K390
 CC K391
 CC K392
 CC K393
 CC K394
 CC K395
 CC K396
 CC K397
 CC K398
 CC K399
 CC K400
 CC K401
 CC K402
 CC K403
 CC K404
 CC K405
 CC K406
 CC K407
 CC K408
 CC K409
 CC K410
 CC K411
 CC K412
 CC K413
 CC K414
 CC K415
 CC K416
 CC K417
 CC K418
 CC K419
 CC K420
 CC K421
 CC K422
 CC K423
 CC K424
 CC K425
 CC K426
 CC K427
 CC K428
 CC K429
 CC K430
 CC K431
 CC K432
 CC K433
 CC K434
 CC K435
 CC K436
 CC K437
 CC K438
 CC K439
 CC K440
 CC K441
 CC K442
 CC K443
 CC K444
 CC K445
 CC K446
 CC K447
 CC K448
 CC K449
 CC K450
 CC K451
 CC K452
 CC K453
 CC K454
 CC K455
 CC K456
 CC K457
 CC K458
 CC K459
 CC K460
 CC K461
 CC K462
 CC K463
 CC K464
 CC K465
 CC K466
 CC K467
 CC K468
 CC K469
 CC K470
 CC K471
 CC K472
 CC K473
 CC K474
 CC K475
 CC K476
 CC K477
 CC K478
 CC K479
 CC K480
 CC K481
 CC K482
 CC K483
 CC K484
 CC K485
 CC K486
 CC K487
 CC K488
 CC K489
 CC K490
 CC K491
 CC K492
 CC K493
 CC K494
 CC K495
 CC K496
 CC K497
 CC K498
 CC K499
 CC K500
 CC K501
 CC K502
 CC K503
 CC K504
 CC K505
 CC K506
 CC K507
 CC K508
 CC K509
 CC K510
 CC K511
 CC K512
 CC K513
 CC K514
 CC K515
 CC K516
 CC K517
 CC K518
 CC K519
 CC K520
 CC K521
 CC K522
 CC K523
 CC K524
 CC K525
 CC K526
 CC K527
 CC K528
 CC K529
 CC K530
 CC K531
 CC K532
 CC K533
 CC K534
 CC K535
 CC K536
 CC K537
 CC K538
 CC K539
 CC K540
 CC K541
 CC K542
 CC K543
 CC K544
 CC K545
 CC K546
 CC K547
 CC K548
 CC K549
 CC K550
 CC K551
 CC K552
 CC K553
 CC K554
 CC K555
 CC K556
 CC K557
 CC K558
 CC K559
 CC K560
 CC K561
 CC K562
 CC K563
 CC K564
 CC K565
 CC K566
 CC K567
 CC K568
 CC K569
 CC K570
 CC K571
 CC K572
 CC K573
 CC K574
 CC K575
 CC K576
 CC K577
 CC K578
 CC K579
 CC K580
 CC K581
 CC K582
 CC K583
 CC K584
 CC K585
 CC K586
 CC K587
 CC K588
 CC K589
 CC K590
 CC K591
 CC K592
 CC K593
 CC K594
 CC K595
 CC K596
 CC K597
 CC K598
 CC K599
 CC K600
 CC K601
 CC K602
 CC K603
 CC K604
 CC K605
 CC K606
 CC K607
 CC K608
 CC K609
 CC K610
 CC K611
 CC K612
 CC K613
 CC K614
 CC K615
 CC K616
 CC K617
 CC K618
 CC K619
 CC K620
 CC K621
 CC K622
 CC K623
 CC K624
 CC K625
 CC K626
 CC K627
 CC K628
 CC K629
 CC K630
 CC K631
 CC K632
 CC K633
 CC K634
 CC K635
 CC K636
 CC K637
 CC K638
 CC K639
 CC K640
 CC K641
 CC K642
 CC K643
 CC K644
 CC K645
 CC K646
 CC K647
 CC K648
 CC K649
 CC K650
 CC K651
 CC K652
 CC K653
 CC K654
 CC K655
 CC K656
 CC K657
 CC K658
 CC K659
 CC K660
 CC K661
 CC K662
 CC K663
 CC K664
 CC K665
 CC K666
 CC K667
 CC K668
 CC K669
 CC K670
 CC K671
 CC K672
 CC K673
 CC K674
 CC K675
 CC K676
 CC K677
 CC K678
 CC K679
 CC K680
 CC K681
 CC K682
 CC K683
 CC K684
 CC K685
 CC K686
 CC K687
 CC K688
 CC K689
 CC K690
 CC K691
 CC K692
 CC K693
 CC K694
 CC K695
 CC K696
 CC K697
 CC K698
 CC K699
 CC K700
 CC K701
 CC K702
 CC K703
 CC K704
 CC K705
 CC K706
 CC K707
 CC K708
 CC K709
 CC K710
 CC K711
 CC K712
 CC K713
 CC K714
 CC K715
 CC K716
 CC K717
 CC K718
 CC K719
 CC K720
 CC K721
 CC K722
 CC K723
 CC K724
 CC K725
 CC K726
 CC K727
 CC K728
 CC K729
 CC K730
 CC K731
 CC K732
 CC K733
 CC K734
 CC K735
 CC K736
 CC K737
 CC K738
 CC K739
 CC K740
 CC K741
 CC K742
 CC K743
 CC K744
 CC K745
 CC K746
 CC K747
 CC K748
 CC K749
 CC K750
 CC K751
 CC K752
 CC K753
 CC K754
 CC K755
 CC K756
 CC K757
 CC K758
 CC K759
 CC K760
 CC K761
 CC K762
 CC K763
 CC K764
 CC K765
 CC K766
 CC K767
 CC K768
 CC K769
 CC K770
 CC K771
 CC K772
 CC K773
 CC K774
 CC K775
 CC K776
 CC K777
 CC K778
 CC K779
 CC K780
 CC K781
 CC K782
 CC K783
 CC K784
 CC K785
 CC K786
 CC K787
 CC K788
 CC K789
 CC K790
 CC K791
 CC K792
 CC K793
 CC K794
 CC K795
 CC K796
 CC K797
 CC K798
 CC K799
 CC K800
 CC K801
 CC K802
 CC K803
 CC K804
 CC K805
 CC K806
 CC K807
 CC K808
 CC K809
 CC K810
 CC K811
 CC K812
 CC K813
 CC K814
 CC K815
 CC K816
 CC K817
 CC K818
 CC K819
 CC K820
 CC K821
 CC K822
 CC K823
 CC K824
 CC K825
 CC K826
 CC K827
 CC K828
 CC K829
 CC K830
 CC K831
 CC K832
 CC K833
 CC K834
 CC K835
 CC K836
 CC K837
 CC K838
 CC K839
 CC K840
 CC K841
 CC K842
 CC K843
 CC K844
 CC K845
 CC K846
 CC K847
 CC K848
 CC K849
 CC K850
 CC K851
 CC K852
 CC K853
 CC K854
 CC K855
 CC K856
 CC K857
 CC K858
 CC K859
 CC K860
 CC K861
 CC K862
 CC K863
 CC K864
 CC K865
 CC K866
 CC K867
 CC K868
 CC K869
 CC K870
 CC K871
 CC K872
 CC K873
 CC K874
 CC K875
 CC K876
 CC K877
 CC K878
 CC K879
 CC K880
 CC K881
 CC K882
 CC K883
 CC K884
 CC K885
 CC K886
 CC K887
 CC K888
 CC K889
 CC K890
 CC K891
 CC K892
 CC K893
 CC K894
 CC K895
 CC K896
 CC K897
 CC K898
 CC K899
 CC K900
 CC K901
 CC K902
 CC K903
 CC K904
 CC K905
 CC K906
 CC K907
 CC K908
 CC K909
 CC K910
 CC K911
 CC K912
 CC K913
 CC K914
 CC K915
 CC K916
 CC K917
 CC K918
 CC K919
 CC K920
 CC K921
 CC K922
 CC K923
 CC K924
 CC K925
 CC K926
 CC K927
 CC K928
 CC K929
 CC K930
 CC K931
 CC K932
 CC K933
 CC K934
 CC K935
 CC K936
 CC K937
 CC K938
 CC K939
 CC K940
 CC K941
 CC K942
 CC K943
 CC K944
 CC K945
 CC K946
 CC K947
 CC K948
 CC K949
 CC K950
 CC K951
 CC K952
 CC K953
 CC K954
 CC K955
 CC K956
 CC K957
 CC K958
 CC K959
 CC K960
 CC K961
 CC K962
 CC K963
 CC K964
 CC K965
 CC K966
 CC K967
 CC K968
 CC K969
 CC K970
 CC K971
 CC K972
 CC K973
 CC K974
 CC K975
 CC K976
 CC K977
 CC K978
 CC K979
 CC K980
 CC K981
 CC K982
 CC K983
 CC K984
 CC K985
 CC K986
 CC K987
 CC K988
 CC K989
 CC K990
 CC K991
 CC K992
 CC K993
 CC K994
 CC K995
 CC K996
 CC K997
 CC K998
 CC K999
 CC K1000
 CC K1001
 CC K1002
 CC K1003
 CC K1004
 CC K1005
 CC K1006
 CC K1007
 CC K1008
 CC K1009
 CC K1010
 CC K1011
 CC K1012
 CC K1013
 CC K1014
 CC K1015
 CC K1016
 CC K1017
 CC K1018
 CC K1019
 CC K1020
 CC K1021
 CC K1022
 CC K1023
 CC K1024
 CC K1025
 CC K1026
 CC K1027
 CC K1028
 CC K1029
 CC K1030
 CC K1031
 CC K1032
 CC K1033
 CC K1034
 CC K1035
 CC K1036
 CC K1037
 CC K1038
 CC K1039
 CC K1040
 CC K1041
 CC K1042
 CC K1043
 CC K1044
 CC K1045
 CC K1046
 CC K1047
 CC K1048
 CC K1049
 CC K1050
 CC K1051
 CC K1052
 CC K1053
 CC K1054
 CC K1055
 CC K1056
 CC K1057
 CC K1058
 CC K1059
 CC K1060
 CC K1061
 CC K1062
 CC K1063
 CC K1064
 CC K1065
 CC K1066
 CC K1067
 CC K1068
 CC K1069
 CC K1070
 CC K1071
 CC K1072
 CC K1073
 CC K1074
 CC K1075
 CC K1076
 CC K1077
 CC K1078
 CC K1079
 CC K1080
 CC K1081
 CC K1082
 CC K1083
 CC K1084
 CC K1085
 CC K1086
 CC K1087
 CC K1088
 CC K1089
 CC K1090
 CC K1091
 CC K1092
 CC K1093
 CC K1094
 CC K1095
 CC K1096
 CC K1097
 CC K1098
 CC K1099
 CC K1100
 CC K1101
 CC K1102
 CC K1103
 CC K1104
 CC K1105
 CC K1106
 CC K1107
 CC K1108
 CC K1109
 CC K1110
 CC K1111
 CC K1112
 CC K1113
 CC K1114
 CC K1115
 CC K1116
 CC K1117
 CC K1118
 CC K1119
 CC K1120
 CC K1121
 CC K1122
 CC K1123
 CC K1124
 CC K1125
 CC K1126
 CC K1127
 CC K1128
 CC K1129
 CC K1130
 CC K1131
 CC K1132
 CC K1133
 CC K1134
 CC K1135
 CC K1136
 CC K1137
 CC K1138
 CC K1139
 CC K1140
 CC K1141
 CC K1142
 CC K1143
 CC K1144
 CC K1145
 CC K1146
 CC K1147
 CC K1148
 CC K1149
 CC K1150
 CC K1151
 CC K1152
 CC K1153
 CC K1154
 CC K1155
 CC K1156
 CC K1157
 CC K1158
 CC K1159
 CC K1160
 CC K1161
 CC K1162
 CC K1163
 CC K1164
 CC K1165
 CC K1166
 CC K1167
 CC K1168
 CC K1169
 CC K1170
 CC K1171
 CC K1172
 CC K1173
 CC K1174
 CC K1175
 CC K1176
 CC K1177
 CC K1178
 CC K1179
 CC K1180
 CC K1181
 CC K1182
 CC K1183
 CC K1184
 CC K1185
 CC K1186
 CC K1187
 CC K1188
 CC K1189
 CC K1190
 CC K1191
 CC K1192
 CC K1193
 CC K1194
 CC K1195
 CC K1196
 CC K1197
 CC K1198
 CC K1199
 CC K1200
 CC K1201
 CC K1202
 CC K1203
 CC K1204
 CC K1205
 CC K1206
 CC K1207
 CC K1208
 CC K1209
 CC K1210
 CC K1211
 CC K1212
 CC K1213
 CC K1214
 CC K1215
 CC K1216
 CC K1217
 CC K1218
 CC K1219
 CC K1220
 CC K1221
 CC K1222
 CC K1223
 CC K1224
 CC K1225
 CC K1226
 CC K1227
 CC K1228
 CC K1229
 CC K1230
 CC K1231
 CC K1232
 CC K1233
 CC K1234
 CC K1235
 CC K1236
 CC K1237
 CC K1238
 CC K1239
 CC K1240
 CC K1

```

Db 709 LRKLKYLSSGVFTGKGVWIPGSESIKIPVCIKVIEDKSGRSQFQAVTDMHMLTIGSLDH 768
Cc -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTRK.
Cc -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
Cc tyrosine phosphate.
Cc -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
Cc (POTENTIAL).
Cc -1- SUBCELLULAR LOCATION: Type I membrane protein.
Cc -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
Cc SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
Cc -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
Cc AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
Cc PHOSPHATIDYLINOSITOL 3-KINASE.
Cc -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
Cc -----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL; U29339; AAC28498.2; -
Cc EMBL; U52530; AAC53050.1; -
Cc HSSP; P13352; 1FGK.
Cc InterPro; IPR000494; EGFR_L_domain.
Cc InterPro; IPR006211; Furin-like.
Cc InterPro; IPR006212; Furin repeat.
Cc InterPro; IPR000719; Prot Kinase.
Cc InterPro; IPR001245; Tyr kinase.
Cc Pfam; PF00757; Furin-like; 1.
Cc Pfam; PF00069; pkinase; 1.
Cc Pfam; PF01030; Recep_L_domain; 2.
Cc PRINTS; PR0109; TYRKINASE.
Cc ProDom; PD000001; Prot_kinase; 1.
Cc SMART; SM00261; FU; 5.
Cc SMART; SM00219; TYRK; 1.
Cc PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
Cc PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
Cc PROSITE; PS50011; MULTIGENE_FAMILY; Receptor; Signal;
Cc Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
Cc Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
Cc SIGNL 1 19
Cc CHAIN 20 1339
Cc DOMAIN 20 643
Cc TRANSMEM 644 662
Cc FT 663 1339
Cc DOMAIN 183 259
Cc FT 707 964
Cc DOMAIN 713 721
Cc NP_BIND 740 740
Cc BINDING 832 832
Cc ACT_SITE 186 194
Cc DISULFID 246 255
Cc FT 202 202
Cc DISULFID 190 190
Cc FT 210 210
Cc DISULFID 214 218
Cc FT 227 235
Cc DISULFID 227 235
Cc FT 231 243
Cc DISULFID 246 255
Cc FT 259 286
Cc DISULFID 259 286
Cc FT 290 301
Cc DISULFID 290 301
Cc FT 305 320
Cc DISULFID 305 320
Cc FT 323 327
Cc DISULFID 323 327
Cc FT 500 509
Cc DISULFID 500 509
Cc FT 504 517
Cc DISULFID 504 517
Cc FT 520 529
Cc DISULFID 520 529
Cc FT 533 549
Cc DISULFID 533 549
Cc FT 536 565
Cc DISULFID 536 565
Cc FT 576 589
Cc DISULFID 576 589
Cc FT 610 621
Cc DISULFID 610 621
Cc FT 617 629
Cc DISULFID 617 629
Cc FT 126 126
Cc CARBOHYD 126 126
Cc FT 250 250
Cc CARBOHYD 250 250

```

```

RESULT 10
ERB3 RAT STANDARD; PRT; 1339 AA.
ID ERB3 RAT
AC 062759; 062955;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-3 precursor (BC 2.7.1.112)
DE (c-erbB3).
GN ERB3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxId=10116;
RX Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
RX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RC MEDLINE=96096535; PubMed=8522190;
RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Koland J.G.;
RT "Cloning of the rat ErbB3 cDNA and characterization of the
RT recombinant protein."
RL Gene 165:279-284 (1995).
RN [2]
RP REVISIONS TO 85; 513 AND 565.
RA Hellyer N.J., Koland J.G.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 922-1097 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Sciatic nerve;
RC MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Froehner P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuroligin and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration."
RL J. Neurosci. 17:1642-1659 (1997).

```

```

PT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 566 566 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
PT CONFLICT 1028 1028 L -> P (IN REF. 3).
SQ SEQUENCE 1339 AA; 147545 MW; 0AA5F2402BBF0F1E CRC64;

Query Match 34.8%; Score 2370.5; DB 1; Length 1339;
Best Local Similarity 41.1%; Pred. No. 6,7e-118;
Matches 527; Conservative 170; Mismatches 431; Indels 155; Gaps 34;

3 LALCRGMLLALLPSPA---STQVCTGDMKRLPASPETHLMDLRHLVYOGQVQGN 59
7 LOYLCC---FLSLARSGEENGQAVCPGTLNGLSTVDADNOYQTLTKLYEKCEVVMGN 62
60 LELTYLPNASLSPLQDIOEVQVYLIAHQVQVPLQRLRIYRGTLFEDNYALAVLDN 119
63 LEIVLGHNAADLFLQWIREVTGYLVANMEFVPLPRLRYVARGTQVQKALFVW-- 120
120 GDEPLANTPTTGASPGSLRLQLRLSTLTKYGLQIRNPOLCYQDTILMKDIFHKNNOL 179
121 ---LANNNT---NSSHALRLQKFTOLTEILSGVYIEKNDKLCMDTIDRDIVRA-- 170
180 ALTLITNRSRACPGSPMCKSGKRCMGESSEDCSLTRTVACAGGC-ARCKPLPTDCHE 238
171 GAETVANKANANCPCHIEVCKG-RCMGPGDDCQILTKTICAPQCNCRCGPPNQCCHD 229
239 QCAAGCTGPRHSDCLACAFHNSGICELHCPALVTYNTDFESNPDEGRYTFGASCTA 298
230 ECAGGCGSGPDTCFACRRFNDGACVRCPEPLVYKTLFQLEPNHTKQYGVGVAS 289
299 CPTNYLSTVGSCTLVCPRLNQBETADGTORCKSKCPKARCVYIGMEHLERAVTS 358
290 CPNPFV-VDDTFCVRACPRDKMEVD-RHGLKMEPCGGLPKACEGSGG-SRYQTVD 345
359 ANIOEPAGCKKIFGSLAFLEPSPGDDPASNTAPLOPBOLOVPELTETITGYLISAPDS 418
346 SNIDGPNCTKLIQNDLFLITGLANDVWHKIPALDEPKLVFRVRIITGLYLIQSPPH 405
419 LPLDSVFQNLQVIRGRILHNGAVS-LTLQGLISWGLRLSLRELKSGALAIHNTILCFV 477
406 MHNFSVSNLTTIGRSLVNRGSLIMKULNTSLGFRSLKELISARAVYISANQOLCYH 465
478 HTVPMDQLFNPQIALHTA-NRPEDECVGBGLACHOLCARHGWGEPPTQCVCNCSQFLR 536
466 HSLMTRRLRGPSEBRRLDIKYDRPLGELAGKVCDDLSSGGCGWGPQCLSCRYNSR 525
537 GQSCVECEVRLQGLPREYVNAHRLPGHPCQOPNGSVTCGPEADOCVA CAHYKDDPFC 596
526 EGVCVTHCNFLQSGPFRFVHAQCFCHPELPMESGSTCGSSGSDCARCAHRRDQPHC 585
597 VAPCGSVKEDLSYMPIMKPEDEGACQPCPINTHSC--VDLDDKCPAQRASPLTSI 654
586 VNSCPHGLD--AKGPIYKYPDAQNECRPHENCTQGCNPELODCLGQAEVLMKSHLV 643
655 ISAVGILLVVLGVVGLIKRQOKIR-KYTRRLQLQETELVEPLTPSGAMENQOMR 713
644 IAVTVG--LAVIIMLIGSGFLYMGRIQNRARARVLEGESEIEPLDPG--EKANKYLAR 700
714 ILKSTELRKVYLSGSGAGTVYVKGIMIPDBENVYKIPALIVLRENTSPKANKELIDAVY 773
701 IFKSTELRKVYLSGSGAGTVYVKGIMIPDBENVYKIPALIVLRENTSPKANKELIDAVY 760
774 MAGVGSFYVRLKGICTLSTVQVLTQAMPYGLLDHRENRGRGLSGDILLNMCQIAKGM 833
761 VGSIDHNIYRLGLCPGSSSLQVLTQVLPGLSLDHKQKRETIQPOLLMNGQIAKGM 820
834 SYLEDVRLVHRDLAARVNLVKSPPNVKXITDPGLARLLDIDETEHADGKVPIMMALES 893

```

```

DB 821 YLREHSWVRDIALANWIKSPSVQVADFGVADLLPDDKOLLHSEAKTPIKMALES 880
QY 1LRRRFTSDVSVSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGERLPORPICTIDVYM 953
DB 881 IHGKXTHQSDVSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGERLPORPICTIDVYM 940
QY 954 IMYKCMIDSECPRRRELVSFSPRNARDPQRFVITQNEMLDGPASPLDSTFYRSLLEDD 1013
DB 941 VYMKCMIDENIPTPEKLEINFTFRMARDPRLVIVKAS-GRQTP--PAAPSVLTTKK 997
QY 1014 MEDLVDAEETLVQGGFCDDPAPGAGVNHHRSSSTSSGGDLTLGLEPSE----- 1068
DB 998 L-----QEALEBEL-----DLPDLAESEGLATS 1023
QY 1069 -----BAPSPAPSEB-----AGSDVPDGLGWAAGLQSLPTHD 1105
DB 1024 LGSALSLPTGTLTPRGSGSLSPSSGYMPMNSISGEALDVAVLGRRQFSPISLH- 1082
QY 1106 PSLQRYSEDPVPLPSETDGYV-----APL-----TC-----SPOPE-----YVQPDV 1145
DB 1083 PIPRGH-----PASESSEGHVTGSEAELOEKVYCSRSRSRSPRPGDSAVHSQHS 1135
QY 1146 RQPPSPRECP-----LPARPGATILRPPTLSP-GKNGVY-----KDYFAF 1187
DB 1136 LTPVTPPLSPGLEEDNGYVMPDTHLRGASSSRSTGLSVLGTBEEDD----- 1191
QY 1188 GSAVENPEYLLPQGAAPQPHPP 1210
DB 1192 -----EEVETMNRKRGRSP-PRPP, 1209

RESULT 11
EGFR DROME STANDARD; PRT; 1426 AA.
ID P0412; O18370; O61601; P81668; Q9WZG0;
AC 13-AUG-1987 (Rel. 05, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.12) (EGfr)
DE (Garten receptor) (Tropo protein) (Drosophila relative of ERBB).
GN EGFR OR TOP OR C-ERBB OR DER OR CS10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId:7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS TYPES I AND II).
RX MEDLINE=94350209; PubMed=8070664;
RA Clifford R., Schubach T.;
RT "Molecular analysis of the Drosophila EGF receptor homolog reveals
RT that several genetically defined classes of alleles cluster in
RT subdomains of the receptor protein.";
RL Genetics 137:531-550 (1994).
RN [2]
RP REVISIONS.
RA Clifford R., Schubach T.;
RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=85124611; PubMed=2982499;
RA Livneh E., Glazer L., Segal D., Schlessinger J., Shilo B.-Z.;
RT "The Drosophila EGF receptor gene homolog: conservation of both
RT hormone binding and kinase domains.";
RL Cell 40:599-607 (1985).
RN [4]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX STRAIN=Oregon-R; TISSUE=Embryo;
RX MEDLINE=87002474; PubMed=3093080;
RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;
RT "Alternative 5' exons and tissue-specific expression of the
RT Drosophila EGF receptor homolog transcripts.";
RL Cell 46:1091-1101 (1986).

```

[5] SEQUENCE FROM N.A. (ISOFORMS TYPE I AND TYPE II), TISSUE SPECIFICITY, AND MUTATION ANALYSIS.
 MEDLINE=99102120; PubMed=9882502;
 RA Leoshkin A.M., Yu S.-Y., Katz J., Baker N.E.;
 RT "Several levels of EGF receptor signaling during photoreceptor specification in wild-type, Ellipse, and null mutant *Drosophila*.";
 RL Dev. Biol. 205:129-144(1999).
 [6]
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I).
 RC STRAIN=Belkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Aghayani A., An H.-J., Andrews-Plamkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bence P.V., Berman B.P., Bhandari D., Bolhakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doull L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durlin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foutel C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodok C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibeagwa C.,
 RA Jaimali M., Kalush F., Karpas G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laevo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Maccari B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshire A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
 RA Spletter R., Spretling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitkies R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [7]
 RP SEQUENCE OF 959-1078 FROM N.A.
 RC STRAIN=Daekwanryeong;
 RX MEDLINE=85137938; PubMed=2983232;
 RA Wadsworth S.C., Vincent W.S. III, Bildeau-Wentworth D.;
 RT "A *Drosophila* genomic sequence with homology to human epidermal growth factor receptor.";
 RL Nature 314:178-180(1985).
 [8]
 RP SEQUENCE OF 1069-1121 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=98401146; PubMed=9731193;
 RA Oates A.C., Wollberg P., Achen M.G., Wilke A.F.;
 RT "Sampling the genomic pool of protein tyrosine kinase genes using the polymerase chain reaction with genomic DNA.";
 RL Biochem. Biophys. Res. Commun. 249:660-667(1998).
 [9]
 RP SEQUENCE OF 1133-1137; 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION ANALYSIS.
 RX MEDLINE=92038942; PubMed=1936959;
 RA Raz E., Schejter E.D., Shilo B.Z.;
 RT "Interallelic complementation among DER/fib alleles: implications for the mechanism of signal transduction by receptor-tyrosine kinases.";

RL Genetics 129:191-201(1991).
 RN [10]
 RP REVIEW.
 RX MEDLINE=97248481; PubMed=9094709;
 RA Peritmon N., Perkins L.A.;
 RT "There must be 50 ways to rule the signal: the case of the *Drosophila* EGF receptor.";
 RL Cell 89:113-116(1997).
 CC -1- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS, WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAF-MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.
 CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES. AND FOR THE DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF ANTIOSEROSA AND VENTRAL NEUROECODERMAL CELLS, GERM BAND RETRACTION, CELL FATE SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF CUTICLE.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Type I;
 CC IsoId=P04412-1; Sequence=Displayed;
 CC Name=Type II;
 CC IsoId=P04412-2; Sequence=VSP_002897;
 CC -1- TISSUE SPECIFICITY: UBIGUOUSLY EXPRESSED IN EMBRYOS. IN LARVAE, UNIFORM EXPRESSION IS SEEN IN WING DISKS. GENITAL DISK. ANLAGEN OF TESTIS AND OVARY. AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW. LEVELS REMAIN HIGH IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX AND THORACIC AND ABDOMINAL GANGLIA.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL, AF053754; AAC08536.1; -;
 DR EMBL, AF053753; AAC08536.1; JOINED.
 DR EMBL, AF053754; AAC08535.1; -;
 DR EMBL, AF053752; AAC08535.1; JOINED.
 DR EMBL, K03054; AAAS1462.1; -;
 DR EMBL, K03417; AAAS1462.1; -;
 DR EMBL, K03416; AAAS0965.1; -;
 DR EMBL, AF109077; AAD26134.1; -;
 DR EMBL, AF109078; AAD26132.1; -;
 DR EMBL, AF109082; AAD26132.1; JOINED.
 DR EMBL, AF109078; AAD26133.1; -;
 DR EMBL, AF109084; AAD26133.1; JOINED.
 DR EMBL, AF109079; AAD26130.1; -;
 DR EMBL, AF109081; AAD26130.1; JOINED.
 DR EMBL, AF109079; AAD26131.1; -;
 DR EMBL, AF109083; AAD26131.1; JOINED.
 DR EMBL, AF109080; AAD26135.1; -;
 DR EMBL, AF003454; AAF46732.1; -;
 DR EMBL, X02293; CAA26157.1; -;
 DR EMBL, AF003912; CAA05747.1; -;
 DR EMBL, X78930; CAA55523.1; -;
 DR EMBL, X78918; CAA55521.1; -;
 DR EMBL, X78919; CAA55522.1; -;
 DR HSSP, P11362; 1FGK.
 DR FLYBASE: FBgn0003731; Egfr.
 DR GO: GO:0007469; P:antennal morphogenesis; NAS.
 DR GO: GO:0003881; P:eggshell pattern formation; IGI.
 DR GO: GO:0007390; P:germ-band shortening; IMP.
 DR GO: GO:0007444; P:imaginal disc development; IMP.

DR GO:0007479; P-leg disc proximal/distal pattern formation; IMP.
 DR GO:0008671; P-maternal determination of dorsal/ventral ax. . .; IMP.
 DR GO:0007477; Pinotum morphogenesis; IMP.
 DR GO:0007314; Pinocyte anterior/posterior axis determination; NAS.
 DR GO:0045468; P-regulation of R8 spacing; NAS.
 DR GO:0016330; P-second mitotic wave (sensu Drosophila); IMP.
 DR GO:0007476; P-wing morphogenesis; IMP.
 DR InterPro: IPR000494; EGFR L domain.
 DR InterPro: IPR006211; Furin-like.
 DR InterPro: IPR006212; Furin repeat.
 DR InterPro: IPR000719; Prot. Kinase.
 DR InterPro: IPR001245; Tyr. Kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.

Query Match 29.0%; Score 1975; DB 1; Length 1426;
 Best Local Similarity 33.0%; Pred. No. 5,7e-97;
 Matches 470; Conservative 166; Mismatches 432; Indels 338; Gaps 41;

24 QVCTGDMKRLRLPASPTLMDLRLHYGCGVVGQVGNLELYLPT-NASLFLQDIDBOVG 82
 100 KICIGTKSRSLVPSNKHNRNLRDRYTNCTYVDGNLKTMLPKNLDSFLDNIREVTG 159
 83 YVLIANOVROVPLQRLRYRTQTLF-----EDNYALAVLDNGDPLNTPVTGASPGSL 137
 160 YLIIHVADYKVVFPRLQIIRGRITLPSLVEEKYALFV-----TYSKM 203
 138 RELQLSLTEILKGVLIQHNPOLCYODTILMKDIFKNNQLALTLIDTRSRACHPCSP 197
 204 YLLEIIDLADLVNGVGFHNHYLCHMRTIQNSEIVSNGTDAYYNDFTAPRECEPCHE 263
 198 MCKGSKWESSEDCSLTFTVAGAGA--RCKGRLPTDCCHQCAAGCTGPRHSDCLAC 255
 264 SCTHG-CWGGGPRXNCKPCKSLTCSPCGAGRCYGPKECHLFCAGGCTGPTQKDIAC 322
 256 LPHNSGICELHPALVTYTDTFESMPNREGRYTPASCTVACPNYVSLTDGSCVLVC 315
 323 KPFDEAVSKEPCPRKKNPTTYVLETNPEGKAYGATCVKSCP--GHLRLDNGACVRSK 361
 316 PLHNGEVTABDGTORCEKSCPCARVYGLGMEHLREAVRANTANIOEPAGCKKIFGSLA 375
 382 PDDKMDKGE-----CVRNGPCPKTCPTGVTVLH-----AGNIDSFNCTVIDGNIR 428
 376 FLPESEDG--DPASNTA-----PLOBQLOVEFTELEITGYLYISAMPDLPDLVSFON 427
 429 ILDTGSGFDVYANTWGPRIYLPDERREVSFTVSEIYGINISBTHQFNLISFRN 468
 428 LOVIRIRILHNGV-SITLOGISWLGSLRELSSGLALIHNTLCTVHTVPMQDLF 466
 489 LETIHROLMESMFALAIYKSSLYSLEMENLKQISSGSVVIQHNRLCYVSNIRMPALQ 548
 487 RNPQALLHTANREDECVGEGLAQHQLCARHGMGPGPTOCNCSQFLAGSCVBEKRV 546
 549 KEPEQVWVWENIRADLCENKGTICSDQCNEDCGWAGTDQCLTKNFENGTCLADCGY 608
 547 LQGLPREYVNAHCLPCHRECPONGSVTCFGEADQCVACAHYKDPFCVACAP----- 601
 609 ISNAVK--PDNRCTCKIHPESCR-----TGNAGADHCQGCYVVRDQHCVSECPKKNYN 660
 602 -----SGVK-----CPI-----NCTH-----PDLSTWPIW 614
 661 DRGVCRCHATCDGCTGPKDTIGIAGCTTGNLAIINNDATVKECLDKDCPD-GY--FW 717
 615 KF--PREBAGACP-----CPI-----NCTH-----632
 718 EYVHPQGSLLKLAGRAVCRKHPCLCELTNYGHEOVSCKTHYRRBQCECECPADH 777
 633 -----SC-----VDLDKXG-----CPAQ 646
 778 YNDBEQRECRQRHPECNGCTGPAADCKSGRNPLPLFANETGPVNSTMNCSTKCLEM 837
 647 R-----ASPLTS-----IISAVVGILLVVVLGVVFGILLKRRQ 679

DB 838 RENVYOTAIQPCASPPRSSKITANLDVNMFIITGAVLVPTICLCV--TYICRQK 895
 QY 680 QKIRKXT--MRLLQTELEVEPLTPSGAMPNQAOMRLKETELRKVYLGSAGFTVYKG 737
 DB 896 QKAKKETVMTALSCGBSEPLRPENIGNANLCKLRIVKQAEIRKGVLGAGGRVYKG 955
 QY 738 IWPDEGNKIPPAIVLRENTSPKANKETLDAAYVAGVSPVSLGLICTSTVQLV 797
 DB 956 VWVPEBENYKIPPAIKELKSTGASSESEPLREAYMASEHVLLKLLAVCMSSQWMLI 1015
 QY 798 TOLMPYGLLDHRENRGLSGODLNMWQIAKMSYLEDVRLVHRDLAARNVLYSPN 857
 DB 1016 TOLMPYGLLDHRENRGLSGODLNMWQIAKMSYLEDVRLVHRDLAARNVLYSPN 1075
 QY 858 HVKIDFGALRLDIDETEHADGKVPFKMMALESILRRRFTQSDVMSYGVTVMEIMT 917
 DB 1076 LVKITPGLAKLSLSSNSNEYKAGGMPIKMLLBCTRNVPFSKSPVMAFGVTIMELT 1135
 QY 918 FGAKPYDGIIPAREIPDLLEKGERLPORPCTIVYIMVCKMIDSECRPRELVSBSFS 977
 DB 1136 FGORPHENIPAKDIPDLIEVGLLEQPEICSLDIYCTLLSCWHLDAAMRPTFKOLLTVFA 1195
 QY 978 RMARDQRFVYIQNEBLS--PASPLDSTFYSLLED--DMGDLVDAEYLVPOQGFPC 1032
 DB 1196 EPARDPGRYALPGDKFTRLPA-----TTSODEKDLIKLAPTTGSEALAKPDYLO 1248
 QY 1033 PDPAGAGGVHRRHSSSTRSGCDLTLGLFSEBEAP-----RSPLAPSEAGSDVF 1066
 DB 1249 PKAPGPS-----HTDCT-----DEMPLKLNKYPKSPKNSSTGDER 1267
 QY 1087 DG--DLGMAKAGLOSLLPHDPSPLORYSEDPVLPSTGDVYALPSCPOPEYVNO 1143
 DB 1288 DSGAREVGGNLR-----LDLPVDEDDYLMPTCPGNNNNNM 1325
 QY 1144 DYRPPSPREBGLPARPARGATLERPKTLSPKNGVYKVPFAGGVENPEYL----TP 1129
 DB 1326 N-----NPNQNNMAVGAAGYM-----DLIGVPVSDNPEYLLAQTL 1364
 QY 1200 QGGAAPQPH-----PPAFSP-AFDNLYYMD 1224
 DB 1365 GVESEPIPTOTIGIPWGGSGTMEVAVPMGSEPTSSDHYND 1408

RESULT 12
 ERBB ALV STANDARD; PRT; 634 AA.
 ID ERBB ALV
 AC P00534;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transforming protein erbB (BC 2.7.1.112).
 GN V-ERBB.
 OS Avian leukosis virus.
 OC Viruses; Retroviral viruses; Retroviridae; Alpharetrovirus.
 OX NCBI_TaxId=11864;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85228222; PubMed=2988784;
 RA Nielsen T.W., Maroney P.A., Goodwin R.G., Rotman F.M.,
 RA Citterione U.B., Raines M.A., Kung H.-J.,
 RT "c-erbB activation in ALV-induced erythroblastosis: novel RNA
 RT processing and promoter insertion result in expression of an
 RT amino-truncated EGF receptor";
 RL Cell 41:719-726 (1985).
 CC -|- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -|- MISCELLANEOUS: THIS PROTEIN IS PRODUCED BY ERYTHROLEUKEMIA CELLS
 CC IN CHICKS WHICH HAVE BEEN INJECTED WITH THE AVIAN LEUKOSIS VIRUS
 CC AND SUBSEQUENTLY DEVELOP LEUKEMIA. THE LEUKEMIAS ARE INDUCED BY
 CC THE INSERTION OF RETROVIRAL PROMOTER ELEMENTS IN OR NEAR THE
 CC C-ERBB GENE, WHICH LEADS TO C-ERBB ACTIVATION.
 CC -|- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ENV-ERBB
 CC PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ldb-eb.ch/announce/>
 CC or send an email to license@ldb-eb.ch).
 CC
 CC EMBL: M10066; AAA48763.1; ALT_INIT.
 CC HSSP: P11362; 1RGK.
 CC InterPro: IPR000719; Prot_kinase.
 CC InterPro: IPR001245; Tyr_kinase.
 CC Pfam: PF00069; pkinase.1.
 CC PRINTS: PR00109; TYRKINASE.
 CC ProDom: PD000001; Prot_kinase.1.
 CC SMART: SM00219; TykC_1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Oncogene;
 CC Glycoprotein; Phosphorylation.
 CC DOMAIN 132 399 PROTEIN_KINASE.
 CC NP_BIND 138 146 ATP (BY SIMILARITY).
 CC FT BINDING 165 165 ATP (BY SIMILARITY).
 CC FT ACT_SITE 257 257 BY SIMILARITY.
 CC SEQUENCE 634 AA; 70891 MW; E705E33A0BB01PCC CRC64;
 Query Match 25.7%; Score 1748.5; DB 1; Length 634;
 Best Local Similarity 52.3%; Pred. No. 2.1e-85;
 Matches 370; Conservative 79; Mismatches 135; Indels 123; Gaps 17;
 QY 587 CAHYKDPFCVACPSGVKPDLSYMPRIKPDSEGAQOPCINQTHSCVLDLDDGCGAEG 646
 DB 3 CAHFIDEPHCVKACPSAGVLSGNDTL-VKTKADANAVQCLCPNCTRGCKGPGLEGCP--- 58
 QY 647 RASPLTSIISAVV-GILLVVVGVVGFILIKRQOKIRKTYMRLLQETLEVEPLPSGA 705
 DB 59 NGSKTPIAGVGVGLLCLVVGIGLILYRR-HYRKQTLRLDLBERLEVEPLTSGE 117
 QY 706 MPQOQMKILKETELRKVKYLGSGAFVTYKGIPIPGENVKIPVAIKLRENTSPANK 765
 DB 118 APQOHLRIKETEFKKYKVLGSGAFVTYKGIPIPGEEKIKPVAIKELRENTSPANK 177
 QY 766 EILDEAVVAGVSPVYSLGICLTSTVOLVQIMPYGCLDHRNRRLSGODLLN 825
 DB 178 EILDEAVVAGVSPVYSLGICLTSTVOLVQIMPYGCLDHRNRRLSGODLLN 237
 QY 826 CMOIAGKMSYLEDLVLRDLAARVAVKSPNHVKTDPGLARLLDIDETEHADGKVP 885
 DB 238 CVOIAGKMSYLEDLVLRDLAARVAVKSPNHVKTDPGLARLLDIDETEHADGKVP 297
 QY 886 IKMALESILRRPETHOSDVMSYGVTVWELMTFGAKPYDGIIPALEIPDLLEKGERLPP 945
 DB 298 IKMALESILRRPETHOSDVMSYGVTVWELMTFGAKPYDGIIPALEIPDLLEKGERLPP 357
 QY 946 ICTIDVYVIMVCKMIDSECRPRRELVSERWAPRQORFVVO-NEDGAPSPLDSTF 1004
 DB 358 ICTIDVYVIMVCKMIDSECRPRRELVSERWAPRQORFVVO-NEDGAPSPLDSTF 417
 QY 1005 YRSLLDDDDMDLVDAEYLVPOGFCPPDPAFGAGGVHHRHSSSTRSGGGLTLGLE 1064
 DB 418 YRSLLDDDDMDLVDAEYLVPOGFCPPDPAFGAGGVHHRHSSSTRSGGGLTLGLE 449
 QY 1065 PSEEBARSPV-----APSGAGSVFDGDLCKMAKGLSLPTHDPSPLORESDDPTVP 1119
 DB 450 -----SRTPLSLSLSATSNNSATNCID-----NGQGHFVREDSFVQRYSSDPTGN 495
 QY 1120 LPSET--DGVAVPITCSPOPEVYNOPVRPQPSRRGRLPAARPAATLERPITLSRGK 1177
 DB 496 FLBSIDGFL-----PAPBYVNO--LMPKGS-----TAMVQ 526

QY 1178 NGVVKOV-----AFGAVENPEYLIFOGGAAPQHPPPAFPAFDNLY 1222
 DB 527 NGVYNNISLTAISKLPMDSRQNSHSTAVDNPEYL-----NTNOSPLAKTVESSPY 578
 QY 1223 WQD-----DPPE-----RGAPSTKGPPTAENPEYLGIDVP 1254
 DB 579 WISGNGHQLNDNPDYQODFLPNETKENGILLKVPAAENPEYLRYAAP 625
 RESULT 13
 ERBB_AVIER STANDARD; PRT; 604 AA.
 ID ERBB_AVIER
 AC P00535;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Tyrosine-protein kinase transforming protein erbb (BC 2.7.1.112).
 GN V-ERBB.
 OS Avian erythroblastosis virus (strain BS4).
 OC Viruses; Retroid viruses; Retroviridae; Avian type C retroviruses.
 OC NCBI_TaxId=79685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H;
 RX MEDLINE=84026539; PubMed=6313229;
 RA Yamamoto T., Nishida T., Miyajima N., Kawai S., Ooi T., Toyoshima K.;
 RT "The erbb gene of avian erythroblastosis virus is a member of the src
 RT gene family";
 RL Cell 35:71-78(1983).
 RN [2]
 RP SEQUENCE OF 1-152 FROM N.A.
 RX MEDLINE=84223957; PubMed=6328658;
 RA Debulire B., Henry C., Benaisa M., Biserte G., Clavierie J.-M.,
 RA Saule S., Martin P., Stehelin D.;
 RT "Sequencing the erbb gene of avian erythroblastosis virus reveals a
 RT new type of oncogene";
 RL Science 224:1456-1459(1984).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- DISEASE: THE V-ERBB ONCOGENE TRANSFORMS AVIAN FIBROBLASTS AND
 CC ERYTHROBLASTS IN CULTURE AND INDUCES SARCOMAS AND ERYTHROLEUKEMIAS
 CC IN CHICKENS.
 CC -1- MISCELLANEOUS: V-ERBB IS A TRUNCATED AND MUTATED VERSION OF THE
 CC RECEPTOR FOR EPIDERMAL GROWTH FACTOR.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ldb-eb.ch/announce/>
 CC or send an email to license@ldb-eb.ch).
 CC
 CC EMBL: X02006; AAA42394.1; ALT_INIT.
 CC EMBL: X01216; AAA42400.1; --
 CC PIR: A06644; TVYU.
 CC HSSP: P11362; 1RGK.
 CC InterPro: IPR000719; Prot_kinase.
 CC InterPro: IPR001245; Tyr_kinase.
 CC Pfam: PF00069; pkinase.1.
 CC PRINTS: PR00109; TYRKINASE.
 CC ProDom: PD000001; Prot_kinase.1.
 CC SMART: SM00219; TykC_1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC Transferrase; Tyrosine-protein kinase; ATP-binding; Oncogene;
 CC Glycoprotein; Phosphorylation.
 CC DOMAIN 132 399 PROTEIN_KINASE.
 CC NP_BIND 138 146 ATP (BY SIMILARITY).
 CC FT BINDING 165 165 ATP (BY SIMILARITY).
 CC FT ACT_SITE 257 257 BY SIMILARITY.

QY 1065 PSEBAPRPL-----APSEAGSDVFDGLGAKGLSLPHTHDSPLQRYSEDPV 1119
 DB 450 -----SRPLSSLSLSTNSATNCIDRNG-----H----- 476
 QY 1120 LPSTGCVAPLTCSPQEPYVQPDVAPPPS 1151
 DB 477 -PVREDGFL-----PAPEYVNG--LMPKKS 499
 RESULT 15
 EGFR CHICK
 ID EGFR CHICK STANDARD; PRT; 703 AA.
 AC P1387;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER
 (fragment)).
 GN EGFR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88261272; PubMed=3260329;
 RA Lax I., Johnson A., Howk R., Sap J., Bellot F., Winkler M.,
 RA "Ulrich A., Vennstrom B., Schlessinger J., Givol D."
 RT "Chicken epidermal growth factor (EGF) receptor: cDNA cloning,
 RT expression in mouse cells, and differential binding of EGF and
 RT transforming growth factor alpha."
 RT Mol. Cell. Biol. 8:1970-1978(1988).
 RL
 CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
 CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
 CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA
 CC synthesis, and cell proliferation.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M20386; AAA48760.1; -
 DR InterPro: IPR000494; EGFRL domain.
 DR InterPro: IPR006211; Purin-like
 DR InterPro: IPR006212; Purin repeat.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; Fu_4_L_domain; 2.
 DR PROSITE; PS00107; PROTEIN KINASE ATP. PARTIAL.
 DR PROSITE; PS00109; PROTEIN KINASE_TYR. PARTIAL.
 DR PROSITE; PS50011; PROTEIN KINASE_DOM. PARTIAL.
 KM Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
 KM Tyrosine-protein kinase; ATP-binding; Phosphorylation.
 FT CHAIN 1 30
 FT DOMAIN 31 >703 EPIDERMAL GROWTH FACTOR RECEPTOR.
 FT TRANSDOM 655 667 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 668 >703 CYTOPLASMIC (POTENTIAL).

FT DISULFID 197 206 BY SIMILARITY.
 FT DISULFID 201 214 BY SIMILARITY.
 FT DISULFID 222 230 BY SIMILARITY.
 FT DISULFID 226 238 BY SIMILARITY.
 FT DISULFID 239 247 BY SIMILARITY.
 FT DISULFID 243 255 BY SIMILARITY.
 FT DISULFID 258 267 BY SIMILARITY.
 FT DISULFID 271 298 BY SIMILARITY.
 FT DISULFID 302 314 BY SIMILARITY.
 FT DISULFID 318 333 BY SIMILARITY.
 FT DISULFID 336 340 BY SIMILARITY.
 FT DISULFID 513 522 BY SIMILARITY.
 FT DISULFID 517 530 BY SIMILARITY.
 FT DISULFID 533 542 BY SIMILARITY.
 FT DISULFID 546 562 BY SIMILARITY.
 FT DISULFID 565 581 BY SIMILARITY.
 FT DISULFID 569 589 BY SIMILARITY.
 FT DISULFID 592 601 BY SIMILARITY.
 FT DISULFID 605 627 BY SIMILARITY.
 FT DISULFID 630 638 BY SIMILARITY.
 FT DISULFID 634 646 BY SIMILARITY.
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 420 420 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 578 578 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT NON TER 703 703 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 703 AA; 77427 MW; APF2DE1B75A650 CRC64;
 Query Match 23.8%; Score 1623; DB 1; Length 703;
 Best Local Similarity 44.7%; Pred. No. 9,9e-79;
 Matches 316; Conservativity 113; Mismatches 250; Indels 28; Gaps 12;
 QY 8 RKGILLALPPGAA-----STGYCTGDMKRLPASPETHMLRLHYCCOVQGL 61
 DB 13 RGAVALVLLLLGALCSAVEBKVCQSTNKLQLQGHVEHFTSLQRYNNCEVLLSLE 72
 QY 62 LLYLPTNASLPIODIOEVGYLLAHNOYROYLRLRYRGTLQFEDNYALVLDNGD 121
 DB 73 ITTVENRDLTKLTKIOEVAGYVLLAHMVDVPLEMLQIIRGVNVLIDNSALVLSYH 132
 QY 122 PLNNTPTVTSAPGRLRLQLRSITLILKGVLLIQRPOLCYODTIIMKDI FHKNNQIAL 181
 DB 133 -NKTQ-----GLRELPMKRLSEILNGSVKISNNPKLCNMDTVLWMDIIDSRSK-PL 182
 QY 182 TLID-TKRSALCPSCMCKGSRCKWSSSDQSLTTRVCAGGA-TCKGLPLPDDCCEQ 239
 DB 183 TVLDFAISNLSSCPKCHNCCTEDHCWAGQONCQLTVITCAQCSGRGKGVPSDDCHNQ 242
 QY 240 CAAGCTGPKASDCLACLFHNSGICELCPALVYNTDTFESMNPBEGRYFGASCYTAC 299
 DB 243 CAGCTGPRSDCLACCKFKPDATCTCPVLYVNTTQMDVNPBEGRYFGATCVAREC 302
 QY 300 PYNLTSDVSSCTLVCPILHNOEYVAEDGTORCEKSKPCARVYGLGMEHLREVAVTS 359
 DB 303 PHNVVTVDHGS CVRS CMTDTYEV-EENGVRKCKKCDGLCSKVCGIGELKGLISINAT 361
 QY 360 NIOFAGCKTFPSGLATLPSFODPASPNTAPLOPBOLOVETLEITGYLYISWPSL 419
 DB 362 NIOFAGCKTFPSGLATLPSFODPASPNTAPLOPBOLOVETLEITGYLYISWPSL 421
 QY 420 PDLVFPNTLQVIRIRILHNGAVSLTQLGLISWGLSLRLSGLALIHNTHLCFVHT 479
 DB 422 TDLVAFENLILIKRTKHOGQSLAVVNLKIQSLGLSLKISISGDIAMNKNLCTVADT 481
 QY 480 VPMDQLFRPHQALLTANRPEDECVGGLAQHQLCARGHGCGPQCVNCSQPLRQOE 539

